

DEMOGRAPHICS OF LYTIC VIRAL INFECTION OF COASTAL OCEAN *VIBRIO*

By
Anne Kathryn Marie Kauffman
B.S., Old Dominion University, 2003

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
and the
WOODS HOLE OCEANOGRAPHIC INSTITUTION

June 2014

©2014 Massachusetts Institute of Technology & Woods Hole Oceanographic Institution
All rights reserved

Signature of Author

Joint Program in Biological Oceanography
Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
and Woods Hole Oceanographic Institution
May 2, 2014

Certified by

Martin Polz
Professor of Civil and Environmental Engineering
Massachusetts Institute of Technology
Thesis Supervisor

Accepted by

Martin Polz
Chair, Joint Committee for Biological Oceanography
Woods Hole Oceanographic Institution

Accepted by

Heidi Nepf
Chair, Departmental Committee for Graduate Students
Massachusetts Institute of Technology

DEMOGRAPHICS OF LYTIC VIRAL INFECTION OF COASTAL OCEAN *VIBRIO*

by

Anne Kathryn Marie Kauffman

Submitted to

the Department of Civil and Environmental Engineering,
Massachusetts Institute of Technology, and the Department of Biology,
Woods Hole Oceanographic Institution,

on May 2, 2014 in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Field of Biological Oceanography

Abstract

Viral predation on bacteria in the ocean liberates carbon from the particulate fraction, where it is accessible to higher trophic levels, and redirects it to the dissolved fraction, where it supports microbial growth. Although viruses are highly abundant in the ocean little is known about how their interactions with bacteria are structured. This challenge arises because the diversity of both bacteria and viruses is exceedingly high and interactions between them are mediated by specific molecular interactions. This thesis uses heterotrophic bacteria of the genus *Vibrio* as a model to quantify virus-host interactions in light of host population structure and ecology. The methods developed in this thesis include streamlining of standard bacteriophage protocols, such as the agar overlay, and facilitate higher throughput in the isolation and characterization of novel environmental virus-host systems. Here, >1300 newly isolated *Vibrio* are assayed for infection by viral predators and susceptibility is found to be common, though total concentrations of predators are highly skewed, with most present at low abundance. The largest phylogenetically-resolved host range cross test available to date is conducted, using 260 viruses and 277 bacterial strains, and highly-specific viruses are found to be prevalent, with nearly half infecting only a single host in the panel. Observations of blocks of multiple viruses with nearly identical infection profiles infecting sets of highly-similar hosts suggest that increases in abundance of particular lineages of bacteria may be important in supporting the replication of highly specific viruses. The identification of highly similar virus genomes deriving from different sampling time points also suggests that interactions for some groups of viruses and hosts may be stable and persisting. Genome sequencing reveals that members of the largest broad host-range viral group recovered in the collection have sequence homology to non-tailed viruses, which have been shown to be dominant in the surface oceans but are underrepresented in culture collections. By integrating host population structure with sequencing of over 250 viral genomes it is found that viral groups are genomically cohesive and that closely-related and co-occurring populations of bacteria are subject to distinct regimes of viral predation.

Thesis Advisor: Martin Polz

Title: Professor of Civil and Environmental Engineering,
Massachusetts Institute of Technology

Thesis Advisor

Professor Martin F. Polz
Department of Civil & Environmental Engineering
Massachusetts Institute of Technology

Thesis Committee

Professor Eric J. Alm
Biological Engineering
Massachusetts Institute of Technology

Professor Sallie W. Chisholm
Department of Civil & Environmental Engineering
Massachusetts Institute of Technology

Professor Janelle R. Thompson
Department of Civil & Environmental Engineering
Massachusetts Institute of Technology

Dr. John B. Waterbury
Department of Biology
Woods Hole Oceanographic Institution

Professor K. Eric Wommack
Department of Plant & Soil Sciences
University of Delaware

Thesis Defense Chair

Professor Edward F. DeLong
Department of Civil & Environmental Engineering
Massachusetts Institute of Technology

ACKNOWLEDGEMENTS

“The way to succeed is to get born at the right time and in the right place.
If you can do that then you are bound to succeed.”

– Sydney Brenner

I was definitely born into the right place, the right time, and all the best company. So many remarkable people have stood behind me, stood beside me, and led me forward - you have brought me strength and joy and lent helping hands in so many things, work and otherwise - I am so thankful to have you with me in this adventure.

I could not have had a better PhD experience than Martin has made possible for me; it has been wonderful in manifold and lasting ways. I am deeply appreciative for the considered balance of independence and support that he has always given me. There is so much wrapped up in being a mentor and over the years I have so often wondered how Martin always seems to get it just right, because he does. The passion, scientific creativity, and kindness of the people Martin brings together in his lab group are a testament to him and make for a community that it is a special privilege to be a part of.

I am thankful to all the members of my thesis committee, Eric Alm, Penny Chisholm, Janelle Thompson, John Waterbury, and Eric Wommack, for the time and thought they have given to my work. Their scientific guidance has helped me to structure my thinking, to avoid and overcome challenges in execution of my work, to learn, and to better understand questions from diverse perspectives. Also, their universal positivity has always left me feeling inspired and invigorated – I hope they realize that I will keep coming back for more.

To my labmates – I love coming to work because of you. Sharing a lab with you has made me a better scientist and a better person; you have made everything better, funnier, smarter, and easier. When I first started I had never streaked bacteria or seen a plaque, you taught me that and so much more with care and kindness. When I started a project that one person couldn't do alone you gave so much of your time and energy even though you had your own work - it simply would not have been possible without you. And all the times I needed help thinking about the questions, the data, and the analyses, you were always ready with time, ideas, information, tools, and scripts. You have shown me the way as scientists, as people, as friends. The number of people who have been in the Polz lab and in the amazing Parsons community over all these years and who have helped me in all my stages of work, and especially in making the sampling possible, make individual acknowledgements a daunting task. I hope that those who know me best can forgive my generality and will understand my exception for one person. Michael Cutler has been the lab manager through my entire time here and he has always made sure that all of us have every last thing we need to do our work, no matter the urgency or the peculiarity of the request. His constant help, advice on all manner of protocols, and collegiality have made a substantial difference in my time here and my ability to achieve my goals.

My family has been a bedrock of support and love for my entire life and I wouldn't be who I am or where I am without them. My mom and dad made me a part of their great adventure, showed me the delights of exploration and discovery, and have always believed in me – it has meant so much. My brother and sister have been my fellow adventurers, they enrich my life and I feel so lucky that we happened to be family. My friends Meredith and Kelley have shown me the meaning of perseverance and strength. And, to my everlovin' and strong husband, Bob - doing science together was best, but Thank You for living through all the moments of every single day of this journey with me, even from 600 miles away. You have lightened my load, helped me be a better scientist, and brought me so much happiness.

I also gratefully acknowledge the WHOI Ocean Ventures Fund, which provided funding for the sequencing of over 250 viral genomes of the Nahant Collection and thereby contributed immensely to the impact of the thesis work presented here. Work presented in this thesis was also made possible by support from National Science Foundation grant DEB 0821391, National Institute of Environmental Health Sciences grant P30-ES002109, the Moore Foundation and the Broad Institute's SPARC program.

TABLE OF CONTENTS

Abstract	3
Acknowledgements	7
List of Figures	11
List of Tables	13
Chapter 1: Introduction	15
Chapter 2: Streamlining Standard Phage Methods for Higher Throughput	47
Chapter 3: The Nahant Collection: Content, Origins, and Nomenclature	73
Chapter 4: Demographics of Lytic Infection in Coastal Ocean Vibrio	115
Chapter 5: Conclusions and Future Work.....	177
Appendix 1: A Simple & Efficient Method for Concentration of Ocean Viruses by Chemical Flocculation	185

LIST OF FIGURES

Chapter 2

Fig. 1	Steps in tubeless agar overlay plating	62
Fig. 2	Images of single plaques generated by streaking-for-singles into molten top agar	63
Fig. 3	Terminase large subunit (LSU) maximum likelihood tree of Nahant Collection phages with terminase LSU proteins from the ACLAMEdb of mobile elements	64
Fig. 4	Recovery of infective virus in plaque forming units over time, following iron flocculation and resuspension in either ascorbate or oxalate solution	65

Chapter 3

Fig. 1	Nahant Time Series sampling site	97
Fig. 2	Three sampling stations in Canoe Cove, Nahant	98
Fig. 3	Example daily datasheet from Nahant Time Series	99
Fig. 4	Sample of images available in Nahant Collection microscopy photo gallery	103
Fig. 5	Phylogenetic tree of Nahant Collection isolates in relation to historical <i>Vibrio</i> collections, based on partial <i>Hsp60</i> sequence	104

Chapter 4

Fig. 1	Phage predators per <i>Vibrio</i> strain as assayed in the large scale direct plating assay, rank abundance curve.....	141
Fig. 2	Host range matrix for assay of phages on <i>Hsp60</i> -sequenced hosts	142
Fig. 3	Phage infection profiles on hosts from all days, with respect to number of host populations and number of hosts infected	144
Fig. 4	Phage genome clusters.....	145
Fig. 5	Electron microscopy of phages.	146
Fig. 6	Phage genomes connected by sharing at least one sequence of ≥ 500 bp at $\geq 99\%$ identity (A), $\geq 95\%$ identity (B), $\geq 75\%$ identity (C).....	155
Fig. 7	Infection profiles of phage genome groups with respect to number of host populations and total number of hosts.	158
Fig. 8	Phage genome group infection profiles, with respect to number of host populations and host sequence types infected.....	161
Fig. 9	Number of infections per <i>Vibrio</i> population, by phage group.	162
Fig. 10	Network representation of host range matrix, with respect to hosts with known populations.	163

LIST OF TABLES

Chapter 2

Table 1	General procedure for tubeless agar overlay	66
Table 2	Comparison of titers in plaque forming units (PFU) as determined by tubeless versus standard agar overlay approaches	67
Table 3	General procedure for serial purification of phage by streaking-for-singles into molten agar overlay.....	68
Table 4	General procedure for archiving and storage of plaques of novel phage isolates.	69
Table 5	Example catalog sheet for tracking of large numbers of plaques	70
Table 6	General procedure for preparation of high titer phage stocks.....	71
Table 7	General procedure for high-throughput spot test for host range.....	72

Chapter 3

Table 1	Locations of data associated with the Nahant Collection	105
Table 2	Nomenclature for Nahant Collection bacterial isolates.	106
Table 3	Catalog of Nahant Collection phages and isolation host	107

Chapter 4

Table 1	Differences in diversity, susceptibility to phages, and phage predator load across days	164
Table 2	Differences in nutrient concentrations across sampling days and over preceding days	165
Table 3	Differences in temperatures across sampling days and over preceding week	166
Table 4	Phage susceptibility of main <i>Vibrio</i> populations	167
Table 5	Distribution of main populations across size fractions, shown as percentage of strains	168
Table 6	Phage infection profile summaries.....	169
Table 7	Properties of Nahant Collection host range matrix & comparison with Moebus & Nattkemper matrix	170
Table 8	Phage group infection profiles	171
Table 9	Summary of expected morphotypes by phage group	173
Table 10	Phage susceptibility of main <i>Vibrio</i> populations, by phage group	175

Chapter One

Introduction

INTRODUCTION

The great abundance of viral particles in the global oceans tells of a hidden network of deadly interactions. Extracellular viral particles are most often borne of host cell lysis and thus their great abundance suggests that viruses play an important role in structuring microbial systems. Though bulk properties of viral production are well described, the structure of the individual transactions that constitute this interaction network are not as well resolved, these are the focus of this work.

In this introduction I provide a historical context for the development of our general knowledge of viral activity in the oceans. The investigations emphasized here are those that are of water column communities, as these most directly pertain to the thesis work presented. I also highlight here that the studies of one particular worker, Karlheinz Moebus, figure especially prominently throughout this introduction as they represent the most extensive studies of the ecology of phage–host interactions ever published and relate intimately to the studies in this thesis. Moebus conducted intensive study of the structure of marine phage-host interactions over two decades beginning in the late 1970s. He began this work after only a few reports of marine phages had been made in the literature and over ten years before their great abundance was demonstrated by transmission electron microscopy (Bergh et al., 1989). To underscore the significance and magnitude of his effort in relation to the general interest of the scientific community at the time it is worth noting some remarks from the conclusion of his first work on the problem (Moebus, 1980). He quotes ZoBell, “Since bacteriophage is generally found associated with large numbers of rapidly multiplying bacteria, it is very doubtful if the sparse bacterial population characteristic of the open ocean is conducive to the development or activity of bacteriophage”. Moebus goes on to say, himself that:

“[The] scarcity of information, at least in part, seems to be due to a general lack of interest in marine bacteriophages as documented by the small overall number of publications dealing with these viruses during the last 3 decades. Admittedly the rareness of phage particles in marine environments poses technical problems which sometimes can be overcome only by considerable effort as demonstrated by the present paper. However, the occurrence of marine bacteriophages indicates that these

viruses are an integral part of their environment and, therefore, should not be neglected (Moebus, 1980)”

First Reports of Marine Viruses

The first report of a confirmed marine phage was by Spencer in 1955, who was working at the Food Investigation Organization and isolated it with the aim of developing a phage for typing of bacterial strains associated with fish spoilage (Spencer, 1955). Spencer used 300mL enrichments containing seawater and multiple marine *Photobacterium* host strains and he reports only that “this method of isolation was repeated many times with various bacterial strains. Bacteriophage was isolated in only one case”, suggesting that any phages present for these strains were present at less than 1 in 300mL¹ (Spencer, 1955). The seawater was collected from 10 miles off shore and the “marine-ness” of the phage was confirmed on the basis of its inability to grow without salt and its sensitivity to heat. Further attempts by Spencer yielded a recovery of 7 phages after investigation of 46L of North Sea seawater using 40 marine heterotrophic bacteria as potential hosts (Spencer, 1960). Of these, 4 phages were recovered by direct plating of unconcentrated seawater and were present at concentrations of less than 10 mL⁻¹, and 3 phages were recovered from 300mL enrichments and present at <0.1 mL⁻¹ (Spencer, 1960).

Further studies also showed that using enrichments, it was possible to recover phages for marine hosts from marine environments. Using a single strain of *Pseudomonas* isolated from the same sample of seawater collected 1 mile off the coast of Chile, Espejo recovered a unique lipid-containing phage in an enrichment with 50mL; total volumes tested are not indicated (Espejo and Canelo, 1968). Baross used 34 previously isolated *Vibrio* hosts and found phages in 2/61 and 5/66 100mL seawater and 10% sediment enrichments, respectively. Using a single previously isolated marine *Vibrio natriegens* (described as *Beneckea* at the time of the study), Zachary screened 140

¹ Spencer also conducted a test for lysogeny in the strains in the mixture giving rise to the single detected bacteriophage to address the possibility that the phage emerged from one strain and proliferated on another, he finds no evidence for lysogeny Spencer, R., 1955. A marine bacteriophage. Nature 175, 690-691.

enrichments using 1-2g mud cores from coastal marshes along the eastern and southeastern United States and found that phages could be isolated from all samples with salinities > 8 ppt (Zachary, 1974).

Abundance of Marine Viruses

Electron microscopy was first used to visualize viruses in 1939 (Kausche et al., 1939), these were filamentous plant viruses. Forty years later Torrella & Morita published electron microscopy (EM) images of viruses in seawater and attached to cells, many of them vibrios (Torrella and Morita, 1979). They prepared their samples by filtering particulate matter onto 0.2um filters and thus could not make general statements about the abundance of viruses free in seawater, but rather, had to limit themselves to estimates of “at least 10^3 viruses per mL in most of the seawater of the bay” though they underscored that any viruses smaller than 0.2um would have been lost and so this was a conservative estimate. They indicate that their study was prompted by the report of Johnson & Sieburth at the ASM Annual Meeting in the previous year indicating that, in Torrella & Morita’s words, “microscopic observations of bacteriophages in offshore waters and bays is not a rare event.” Unfortunately, these authors did not follow up their work with studies on unfiltered samples. Thus, it was another 10 years before the report came that viruses could be present at concentrations of up 10^7 mL^{-1} in the ocean, also demonstrated by EM (Bergh et al., 1989). Curiously, Torrella and Morita’s work were not cited in the 1989 paper, underscoring how little impact these early observations seem to have made on the field. Increasing reports of the great abundance and activity of bacterial cells in the ocean, belied by culture based studies, were likely the necessary precursor to great interest in the phages that might infect them (Cho and Azam, 1988; Hobbie et al., 1977; Jannasch and Jones, 1959; Sherr, 1989; Watson et al., 1977).

Viral Production

Widespread embrace of the view of the ocean as a busy milieu of 10^6 mL^{-1} bacteria and 10^7 mL^{-1} viruses stimulated a decade of research rich in studies

characterizing the bulk dynamics of viral production in the oceans (Wommack and Colwell, 2000). Among the first attempts to address the impact of viral infection on bacteria was transmission EM study of the numbers of visibly infected cells (Proctor and Fuhrman, 1990). On average ~1-4% of heterotrophic bacteria and ~0.8-2.8% of cyanobacteria were observed to contain viral particles (Proctor and Fuhrman, 1990). Estimates of the total number of cells infected based on these numbers require a conversion factor based on the percentage of the latent period during which viral particles are visible. Using a *Cytophaga marinoflava* model in which viral particles are only visible in the majority of cells in the final 10% of the latent period, the authors estimated that ~10-40% and 8-28% of heterotrophic bacteria and cyanobacteria, respectively, were infected (Proctor and Fuhrman, 1990; Valentine and Chapman, 1966). Assuming steady state conditions wherein bacterial production is equivalent to loss, these yielded estimates of 60% and 30% virally-mediated mortality for the two groups. As pointed out by Waterbury, however, estimates of the percent of latent period with visible viruses is at least 50% for cyanoviruses, suggesting that far fewer cyanobacteria are infected or killed by viruses than suggested (Waterbury and Valois, 1993). Nonetheless, this provided evidence for active infection of bacteria by phages in the ocean.

Characterizations of the abundances of viruses and bacteria in diverse aquatic environments were carried out using electron and epifluorescence microscopy [compiled and reviewed in (Wommack and Colwell, 2000)]. In aggregate, these studies showed viruses to be abundant, generally correlated with bacterial abundance, and abundant when bacterial productivity is high, and variable in composition, as assessed by viral morphology (Wommack and Colwell, 2000). Decay of viral infectivity was shown to be more rapid than physical loss, and accelerated by exposure to sunlight. Decay rates were also shown to vary across viral types, sometimes with evidence for adaptation to local levels of irradiance such that, for example, viruses isolated from higher latitudes decayed more rapidly at lower latitudes than did locally isolated viruses (Noble and Fuhrman, 1997; Wilhelm et al., 1998; Wommack et al., 1996). These studies also showed differences across environments, with decay rates lower in oligotrophic oceans (Wilhelm

et al., 1998). Evidence for differences based on the treatment of water samples with autoclaving, filtration, or metabolic inhibitors, suggested that the presence of living cells and particulate matter directly impacted the decay rates of viruses and providing a mechanism for explaining variability in observed estimates. On average, decay rates of viable marine phages in natural seawater can be approximated at about $\sim 2.7\% \text{ h}^{-1}$ [based on compiled studies in (Wommack and Colwell, 2000) and omitting few outliers with $>30\%$ decay rate per hour].

Composition of viruses is variable across time and space. Studies that demonstrate this have included community level genetic fingerprinting as well as culture based studies of interactions with specific hosts, a few studies of the former are highlighted here and the latter are discussed in the context of host-specific interactions. In a study of the Chesapeake Bay virioplankton using pulsed field gel electrophoresis (PFGE) it was shown that fingerprints based on presence, absence, and relative abundance of different viral genome size classes, clustered by sampling date and geographic region. This demonstrated that viral communities were dynamic and, given the assumed specificity of interactions, likely to impact bacterial communities differentially based on their changing composition (Wommack et al., 1999). In a study of *Synechococcus* phages, also in the Chesapeake Bay, a conserved gene for assembly of the viral capsid (gp20) was used as a target for terminal restriction fragment length polymorphism (T-RFLP) fingerprinting and viral diversity was tracked at a single station at 9 time points over a year (Wang et al., 2011). Fingerprints had features of both variability and stability, with changes in composition evident through time but also with specific peaks steady over several consecutive months before receding in abundance. In an attempt to generate a community-wide interaction network a three year time series of monthly resolution was carried out at the San Pedro Ocean Time series (SPOT); operational taxonomic units (OTUs) for different groups were tracked, using automated ribosomal intergenic spacer analysis (ARISA) for bacteria and TRFLP of the conserved T4-like myovirus major capsid protein (gp23) (Chow et al., 2013). It was found that there were several patterns of abundance of viral TRFLP OTUs, with some showing stability over the entirety of the

sampling period and also some viral OTUs correlating with either single or multiple bacterial OTUs at the resolution of 1 month used in this study. However, as single TRFLP OTUs subsume extensive genomic diversity it is not clear to what extent this reflects true stability at the level of closely related viral genotypes. Shifts in viral assemblage composition at the level of TRFLP OTUs were more predictable than those in bacterial assemblages and were related to day length, season, salinity, and temperature but these differences may reflect differing extents of diversity encompassed by viral and bacterial markers. Another study was conducted using the same approach but with samples collected at daily resolution for 38 days and near Santa Catalina Island, on the US West Coast (Needham et al., 2013). It was found that “common bacterial and viral taxa were consistently dominant, and relatively few displayed dramatic increases/decreases or ‘boom/bust’ patterns that might be expected from dynamic predator/prey interactions.” Significant correlations were, however, observed between viral and bacterial OTUs and these occurred with a median 2 day time lag, interestingly, it is noted by the authors that this is shorter than that predicted based on predictions of marine bacterial growth rates and suggests that it is disproportionately faster growing hosts that are impacted by viral infection. These studies all suggest that there is some stability in association and interactions between viral and bacterial at some level of taxonomic resolution and that environmental parameters play a role in defining members of interacting communities.

Impact of Operational Choices on Studies of Viruses

Several considerations suggest that our understanding of bulk viral accounting may yet change. The elements that we now describe as viruses were originally described as “filterable viruses”, at the time of this usage the term “virus” simply connoted a poisonous agent and, whereas many agents of disease were being identified as bacteria and retainable on a filter, there were some disease causing (“poisonous”) agents that could not be removed by filtration [reviewed in (Hughes, 1977)] . Filtration has continued to be the most common operational method for separating viruses from cells.

However, there is overlap in size classes between viruses and cells and thus any approach that uses filtration to remove cells can result in loss of large viruses, as well as in classification of non-viral elements as viruses. The discovery of viruses that are nearly 1µm in size was a surprising finding in part simply because filtration as a common step had precluded their discovery; it is now becoming apparent that such viruses may be common – though thus far they are specific to eukaryotes not bacteria or archaea (Bratbak et al., 1990; Claverie and Abergel, 2010; Philippe et al., 2013). Filtration also necessarily removes larger particulates, which have been shown to have associated phages and may be significant sites of viral replication given the abundance of bacteria and the access of those hosts to nutrients and thus their more rapid growth (Bratbak et al., 1990; Proctor, 1991; Torrella and Morita, 1979). Treatment of samples with chloroform to kill bacteria instead of removing them is effective and offers one alternative, though it damages lipid containing viruses (Fredericq, 1950). Ahrens compared the recoveries of viruses from samples either treated with chloroform or pre-filtered to remove bacteria, she found that recoveries of plaque forming units (PFU) was highest with chloroform treatment and declined with decreasing pore size of pre-filters, as follows: chloroform (780 PFU mL⁻¹), 0.6µm filters (140 PFU mL⁻¹), 0.45µm filters (94 PFU mL⁻¹), and 0.2µm filters (37 PFU mL⁻¹) (Ahrens, 1971). Importantly, the use of filtration can also lead to the misclassification of a variety of entities as viruses. Several non-viral biological elements can pass through 0.2µm filters and be misclassified as viruses, these include small cells, gene transfer agents (GTAs), and membrane vesicles (Forterre et al., 2012; Mendes et al., 2013). GTAs are known to be present in marine bacteria and capable of mediating transduction between them, though they are presently thought to be rare based on homology-based searches dependent on recognition by similarity to known GTAs (Biers et al., 2008; McDaniel et al., 2010). Depending on how viruses are detected in steps these different elements can impact judgments of abundance and viral sequence composition. In the case of epifluorescence microscopy, cellular, GTA or vesicle contamination will give false positives, in the case of electron microscopy GTAs will give false positives. Also, many true viruses, particularly RNA and ssDNA viruses, are

missed because they are not easily detected using standard epifluorescence microscopy approaches for viral enumeration (Holmfeldt et al., 2012).

Importantly, membrane vesicles have recently been shown to occur at concentrations of up to $6 \times 10^6 \text{ mL}^{-1}$ in the ocean (Biller et al., 2014). Given that concentrations of viruses in coastal oceans are often approximately $\sim 10^7 \text{ mL}^{-1}$ it is reasonable to consider that epifluorescence count-based studies could be substantially confounded by contribution of such vesicles if their production is also variable in time. Further, the observation that viruses appear to inject their DNA into these vesicles, and that phage sequences are detectable in environmental vesicles, adds an entirely novel reservoir for “live” viral genomes that has not previously been considered (Biller et al., 2014). Such vesicular reservoirs would act as sinks with respect to detectable infective viruses though they would still be resident in the community and potentially infective, as for example, membrane vesicles have been shown to mediate transfer, and subsequent expression, of genes between bacteria (Yaron et al., 2000). Transfer to such vesicles also offers the potential for eluding receptor-based requirements for entry and thus create a disjuncture between the measurable and actual host range of such viruses when investigated in their encapsidated form. In any case, it is likely that the dynamics of production of such vesicles plays a role in the interplay between hosts and phages and these processes have thus far either been integrated into current bulk estimates or missed. All current approaches for studying viruses are thus admittedly not wholly inclusive nor entirely exclusive of non-viral elements.

Host-specific Phage Predation

Description of the structure and dynamics of predator-prey interactions in a community requires fundamentally that the number of interacting members – predator and prey – be known. For phage-host interactions quantitative estimates of the number of predators and prey present in the same sample are limited. Although bulk counts of bacteria at the population level bound estimates of maximum total numbers of interacting members they cannot resolve the true number of interactions, which occur between

compatible strains of highly diverse phages and bacteria. The great genotypic diversity of bacteria in natural environments and the generally low abundance of specific phage types pose logistical challenges to achieving such estimates. Attempts to address some component of the question of strain-level susceptibilities of hosts to phages *in situ* were made in particular by the studies of Ahrens (Ahrens, 1971), Hidaka (Hidaka, 1977), Moebus (Moebus, 1980), and Waterbury (Waterbury and Valois, 1993), and these are presented in detail in the following sections. Though tests of host range associated with these studies are mentioned here briefly they are explored in greater detail in later sections.

In 1971, Ahrens carried out quantitative surveys of the relative abundance of *Agrobacterium* (later reclassified to *Stappia*, *Ruegeria*, and *Ahrensia* species) and *Agrobacterium*-specific phages with respect to salinity, in the brackish Bay of Kiel, between the North and Baltic Seas (Ahrens, 1971; Uchino et al., 1998). She emphasized that the Baltic Sea afforded a unique location to study the distributions and dynamics of host-specific phages because *Agrobacterium* were present at concentrations of up to several thousand per mL and their phages present in concentrations high enough to detect by direct plating without enrichment or concentration. She used the chloroform method of Fredericq (Fredericq, 1950) to preserve phage samples without the need for filtration and thus facilitate her collection of larger numbers of samples without bottlenecks for processing by filtration. She collected samples at 27 stations, in some cases at multiple depths, and at one station in a time series of 14 samples over 6 months. She assayed for both coliphages and *Agrobacterium*-phages and showed that whereas the phages of the former rapidly declined in titer at distance from freshwater sources, the *Agrobacterium*-phages declined below salinities of less than 8ppt. Using two previously isolated indicator strains, both *Agrobacterium stellulatum* (now *Stappia*), she showed that phage concentrations on each of the indicators were different and dynamic in time. The highest concentrations observed were of 36,500 PFU mL⁻¹, with concentrations on the same indicator <2,000 mL⁻¹ at the next test 2 weeks later. Decay rates assayed in natural brackish water in the lab showed that, for one phage, concentrations dropped from 10⁵

mL⁻¹ to undetectable after 4 weeks, and for another that they dropped from 2*10⁶ mL⁻¹ to <1 PFU mL⁻¹ in 2 weeks. The abundances of phages and bacteria in a time series at a single station followed similar trends over 6 months, with phage abundance (as assayed by specific indicator strains) generally substantially lower than host group abundance. Ahrens concluded that the abundance of phages was dependent on bacterial growth, with short periods of intensive growth followed by equally short periods of die-off, though it was not clear that these should be attributed to phages rather than changes in hydrographic conditions. Prior to her study of *Agrobacterium* phages in the Bay of Kiel, Ahrens conducted taxonomic investigations of the genus and showed that physiologic properties were homogeneous within *A. stellulatum* and *A. ferrugineum*, respectively, but different between them (Ahrens, 1968). For this reason, she posed the question of whether phage host ranges could also be seen to discriminate these species. To address this she conducted a host range cross test with 39 phages isolated from several species (*A. stellulatum*, *A. ferrugineum*, *Agrobacterium* strain C4, *E. coli*, and *A. radiobacter*) against 53 previously isolated hosts of these same species and also a *Caulobacter*. She found 60/2067 positive interactions, with 19/39 infecting only their host of isolation and no phages infecting multiple species, though some did infect multiple strains within a species. On the basis of these studies Ahrens concluded, as future workers doing similar studies would also, that the diversity of cells within taxonomic groups was exceedingly high and that encounter rates between any one specific bacterial lineage and its specific phages must generally be exceedingly low such that “despite high numbers of phages the agrobacteria are never fully decimated.”²

During a cruise as part of a “marine ecology survey” in the southwestern Pacific Hidaka addressed the question of the prevalence of host co-existence with predatory phages, also with respect to host distributions and abundance (Hidaka, 1977). Samples were collected at 13 stations along a transect and at 5 depths down to 300m. Bacteria in these samples were enumerated by the most probable number method and by plating, and colonies developing on the count plates were isolated for further study. Seawater samples

² In-text quote translated from German.

were also collected at each of the stations and depths for the study of phages. Phage samples were amended with nutrients, as well as with “culture” (unfortunately it is not clear what this culture was of) and incubated overnight before storage in a cold room for later processing and study. A collection of 576 bacterial isolates were assayed for susceptibility to phages deriving from enrichments of the same sample and 72 (12.5%) were found to be sensitive. Attempts to recover phages using direct plating were generally not successful, indicating concentrations of phages were likely less than at least 1 in 10mL³. Heterotrophic bacteria were most abundant at the surface and 50m collection depths and the recovery of phage-host systems correlated with this abundance of bacteria, with many of the deeper depths not yielding phage-susceptible bacteria (though bacterial isolates were also fewer at these depths). These results suggested that the abundance of bacteriophages correlated generally with the abundance of bacteria. The phage hosts were identified as belonging to the *Pseudomonas*, *Vibrio*, *Photobacterium*, and *Lucidobacterium* (later, *Vibrio harveyi*). To develop a sense of the abundance of specific phage hosts it is worth considering the following examples: 1) at the station with the maximum concentration of colony forming units (CFU), 110 CFU mL⁻¹, at least 3 independent phage-host systems were recovered, suggesting that each host genotype was likely present at not greater than 30 cells mL⁻¹, 2) at a station with 20CFU mL⁻¹, 4 phage-host systems were recovered, suggesting that each host genotype was likely not present at greater than 5 cells mL⁻¹. These results are remarkably consistent with much later sequence-based demonstration of the exceedingly low abundance of specific genotypes present in seawater (Thompson et al., 2005). Given that enrichments were successful, but that direct plating was not, this suggests that phages were present at 0.003 – 0.1 mL⁻¹, more than a magnitude less abundant than their specific host genotypes as estimated above.

³Specific methods used in relation to treatment of phage samples are unclear on several counts but it is patent that the intent was isolation of phages from the same sample as the bacteria: “The marine bacteriophages in the sea water samples were detected by lytic action for the bacteria isolated from the same samples.” Methods and volumes used to attempt direct plating are not indicated but this liberal estimate is based on the use 10mL volumes as described by Spencer (Spencer, R., 1960. Indigenous marine bacteriophages. *Journal of Bacteriology* 79, 614.

Moebus subsequently also conducted a large scale study of the ability to co-isolate phages and hosts from the same sample. He too isolated bacteria from many ocean transect stations but he used nutrient enrichment to support growth of native bacteria and thus increase viral concentrations. Over an 11 week cruise from the European continental shelf to the Sargasso Sea and back, he collected samples at 48 stations. From each of these samples he isolated 30 bacterial strains immediately after collection. To collect phages he prepared three subsamples of water from the same sample used to isolate bacteria and amended these with peptone and yeast extract to stimulate native bacterial growth and increase viral production. The three subsamples were amended after different hold times, either immediately, after 24 hours, or after 48 hours, and phage samples were collected from each of these 1, 2, and 3 days post-amendment. Using this approach Moebus collected 1,382 bacterial isolates, of these only 913 survived to be re-cultured. He screened each of the surviving 913 strains and found 213 (23%) to be susceptible to phages from the same sample (Moebus, 1980). Subsequent investigations of the taxonomy of phage-sensitive isolates showed the entire collection of sensitive strains to be in the family Vibrionaceae, with ~77% in the genus *Vibrio* (Moebus and Nattkemper, 1983). Investigation of the taxonomy of insensitive strains led to the conclusion that while these were more diverse at least 37% of all original isolates were in the family Vibrionaceae. Moebus's conclusion based on this determination of taxonomy was that vibrios (and their phages) were likely abundant in the Atlantic Ocean but also that they must be preferentially selected for under nutrient enrichment such that vibriophages come to dominate bulk viral production under such conditions (Moebus and Nattkemper, 1983).

To address the abundance of phages *in situ*, Moebus conducted a subsequent study using indicator strains in direct plating assays of seawater to determine the concentration of plaque forming units (PFUs). He conducted a time series in which he enumerated PFUs in 35 samples spanning 6 months. He used at least 35 indicator strains isolated the previous year, and isolated additional indicator strains during the course of his time series to add to his collection of indicators. Using these indicators he found that

phage concentrations of 0 to 2 PFU mL⁻¹ dominated but that concentrations with some indicators could briefly reach 1,500 mL⁻¹. His observations revealed that phage composition varied over time as different indicators showed different distributions of phage concentration over time. This study also showed that spikes in phage production detected by a given indicator could decay rapidly, with observed decay rates ranging from 0.3 to 0.92 day⁻¹ (Moebus, 1992). These investigations showed that the presence of specific phages in the environment in great abundance was transient and that, at least with indicator strains, concentrations of phage predators were usually low to undetectable. Regarding the maintenance of phage host systems given his observations, Moebus concluded that:

“The concentration of suitable host cells, as well as that of many lines of matching virions, is always low and may decrease, at least during winter, to levels too low to warrant timely phage reproduction that would prevent extinction of the PHS [phage host system] in question. For these considerations it seems plausible to assume that maintenance of PHS depends mainly on phage survival within cells. Essentially, there are three qualitatively different possibilities: lysogeny, the carrier state, and delayed reproduction or release” (Moebus, 1992).

Extending these findings that predator loads at the strain level were likely low was an elegant study by Waterbury that used dilution approaches to distinguish susceptibility of host strains based on their relative abundance. This study found that the most abundant *Synechococcus* strains, those present in the highest dilutions of source seawater, were resistant to co-isolated phages, whereas those present at the lowest concentrations were susceptible. Further, the recovery of different strains in the enrichments of less-diluted samples than in those of more-diluted samples suggested that the susceptible strains present in low abundance were also better resource competitors. These results showed that the majority of *Synechococcus* cells in the environment were resistant to phages and that these cells were less fit in resource competition than the rare phage-susceptible cells. In this same study, a two-year, monthly-resolution, time series study was carried out to determine the relative abundance of *Synechococcus* and their

phages and to assess whether declines in host abundance were associated with phage production. Highly sensitive indicator strains (isolated in previous studies) were used to provide estimates of total *Synechococcus* phage abundance. This time series study showed that overall *Synechococcus* phage abundance generally mirrored total *Synechococcus* abundance, though at an order of magnitude lower concentration, and that host abundance was predictable based on season and was independent of phage concentrations. Estimates of the mortality associated with phage predation, based on the observed numbers of total *Synechococcus* and associated phages, ranged from 0.005-3.2%. As indicated by the authors, these numbers did not account for the finding that the host assemblages were dominated by resistant strains and thus that encounter rates should be significantly lower than those projected based on total concentrations. These findings suggested that top-down control of *Synechococcus* by their phages was negligible.

A subsequent study by Hennes and Suttle explicitly addressed the capacity of naturally occurring phages to control blooms of specific host genotypes (Hennes et al., 1995). By fluorescently labeling viruses specific to a particular host genotype, and using them as probes, they were able to enumerate the abundance of that strain over time in a mixed community. Using this approach they showed that amendments of large numbers of a specific host genotype to natural seawater samples elicited emergence of viral predators that increased from undetectable to nearly 73% of the total viral abundance in the case of nutrient amended treatments (Hennes et al., 1995). Replication of these viruses drove down the relative abundance of the spiked-in host strain from 43% to 0.01% of the total community within 12 hours in nutrient amended treatments and from ~30% within 32 hours in unamended treatments.

Finally, in the only such study known to this author, the abundance of a very narrowly defined class of phage hosts was tracked daily over 3 months in parallel with direct plating assays for abundance of plaque forming units (PFUs) specific to a previously isolated host in the same class (Jost and Wiese, 2013). Polyclonal antibody markers were stimulated for this historical strain, *Sphingomonas* B18, and then used to track *in situ* abundances of strains susceptible to these same antibodies over a period of

three months. All colonies with *Sphingomonas* morphology recovered for each day were tested for reaction with the antibody as well as for susceptibility to two control phages previously isolated and infective on B18. In this way, total CFU and antibody/control-phage positive CFUs were enumerated. Daily abundance of PFUs was then determined quantitatively using agar overlays of seawater with B18 as the test strain. Diversity among antibody-reactive strains did exist, as evidenced by the differential susceptibility of some strains to infection by the two control phages. The CFU of these narrowly defined hosts, and the PFUs determined using the highly similar indicator strain, were used to estimate encounter rates using the commonly used formulation of Murray and Jackson (which takes into account the Sherwood number, cell diameter, viral diffusivity, and cell and phage concentrations) and these were found to be too low to sustain the observed viral production (Murray and Jackson, 1992). Next, and on the basis of previously described microscale patchiness of organism and organic materials in the marine environment (Long and Azam, 2001; Stocker, 2012; Stocker et al., 2008; Verdugo et al., 2004), they calculate a concentration factor and minimum volume within which phages and cells would have to co-occur to allow for sufficient interactions to give rise to observed changes in PFU. In this way they determined that if all the interacting members were present in volumes of 234 and 37uL (for a required 1 or 6 interactions $\text{mL}^{-1} \text{ day}^{-1}$ respectively) that sufficient encounters would occur to sustain observed production. These observations suggest that small-scale spatial patchiness in phage and host concentrations, perhaps resulting from association with particles or dense chemotactic aggregations, may provide sufficient local encounter rates to sustain phages that otherwise seem unlikely to encounter host before they decay.

All of these studies have provided large scale overviews of some aspect of specific phage-host interactions *in situ*. The authors are unanimous in their conclusion that the diversity of both hosts and phages is very high, and that the specific encounter rates between most compatible pairs must thus generally be exceedingly low. The use of different indicator strains on the same samples had shown changing abundances of phages with a given specificity, suggesting that environmental hosts susceptible to these

classes of phages would also be subject to cycles and thus that strain-level bacterioplankton composition was likely changing as a result of phage infection. Together with the demonstration, albeit using indicator strains, that phage abundance was generally lower than overall host group abundance, it appeared that phage predation was not likely a significant agent of top down control for the majority of strains in the environment at any one moment in time, though it was likely to impact strain level composition. Workers conducting phage susceptibility assays on environmental bacteria were therefore likely the first to appreciate that extensive genotypic diversity did exist among otherwise indistinguishable strains, and that it was manifest specifically in relation to susceptibility to a potentially significant class of predators.

Modeling Phage-Host Interactions

A model was developed that integrated these emerging findings to set forth a theoretical framework allowing for the coexistence of diverse host strains as a consequence of viral predation (Thingstad and Lignell, 1997). By assuming exclusive predator-prey bacterial-host pairs in a nutrient-limited, and non-selective grazer controlled system, this model allowed for steady state stable coexistence of multiple host types. In this model the fastest growing hosts sustain the largest numbers of viruses and their suppression by these viral predators leaves resource available allowing for invasion and establishment of inferior competitors. Phage predation thus effectively “kills the winner” (KtW) and prevents competitive exclusion. Incorporation of a cost of resistance (COR) yields inverse rank abundances of host types and their associated viruses, such that host types with the lowest COR, the competition specialists, are the least abundant and their viruses the most abundant (Våge et al., 2013a). The discovery of phages for the most abundant bacterial group in the ocean, the SAR11 *Pelagibacter* clade, was interpreted by some to indicate that KtW was wrong because under the model this dominant bacterial group should be highly phage resistant (Zhao et al., 2013). However, by assuming the COR is distributed across genotypes within a population rather than across higher taxonomic units, KtW still applies (Våge et al., 2013b). In this case, it is

predicted that there are rapidly growing sensitive strains and slowly growing resistant strains within every population, including SAR11. This also exactly describes the structure of interactions documented for the *Synechococcus* (Waterbury and Valois, 1993), wherein the most abundant strains were also more resistant and less fit resource competitors than susceptible rare strains. These results suggest that, despite the diminutive 1.3Mb genome, there is likely microdiversity within the SAR11 *Pelagibacter* and studies of genome diversity and rates of recombination support this expectation (Acinas et al., 2004; Gilbert et al., 2008; Giovannoni et al., 2005; Rusch et al., 2007; Vergin et al., 2007; Wilhelm et al., 2007).

Microdiversity, however, itself poses a challenge for our understanding of phage-host interactions as guided by KtW. Extensive lateral gene transfer among microbes shifts the impact of negative frequency dependent selection from genotypes to individual genes. As a consequence, phage infections dependent on the presence or absence of specific genes impact a crowd of microdiverse genotypes rather than a single lineage. Microbial populations in the environment are units of gene flow that display ecological cohesion (Cordero and Polz, 2014). Within these microbial populations homologous recombination can be extensive and frequent such that nucleotide identity across the core genome can be very high but average shared gene content can be as low as 70% (Cordero and Polz, 2014). Low frequency phage receptors are overrepresented in the mobile gene pool and are thus present in diverse genotypes within a population. This creates a scenario whereby phage predation acts in a negative frequency dependent manner on single phage-receptor genes rather than on whole genome types. Given such a scenario we expect that phages requiring the presence of a specific receptor allele may be found to have multiple hosts that are distinct by some other marker used to assess phylogeny, though they are identical at the requisite receptor. This perspective underscores the need for distillation of the concept of phage host-range to explicitly address whether it is a function of the host-range of the receptor or the capacity of the phage genome to replicate in divergent genomic backgrounds. This view also suggests that while negative frequency dependent selection may in fact contribute substantially to viral production it may not be

able to prevent competitive exclusion by certain core genotypes provided they expand under a cloak of highly recombining and low-frequency receptor diversity.

Phage host range also poses a challenge for KtW in that this model rests on the assumption of exclusive pair-wise interactions. However, a multi-host infection strategy would effectively increase encounter rates and thus be expected to provide a selective advantage to phages that were capable of it. Whereas single-host phages are projected to provide a mechanism for maintaining diversity, multi-host phages could act to diminish diversity in this model by eliminating low abundance genotypes when predators are subsidized by increased growth rates of alternative hosts. Indeed, the most recent iteration of KtW does incorporate multi-host phage infection profiles, wherein each phage infects a hierarchically nested subset of all hosts (Vage et al., 2013). This model predicts that generalist phages have lower adsorption constants and are less abundant than highly specific phages, which have higher adsorption constants. However, as described in the following sections, none of the phage and host collections used to demonstrate host range meet both the criteria of: 1) sampling hosts and phages at the same time, and 2) isolating these phages in a quantitative way (rather than using enrichments). Therefore, the prevalence of phages with different host ranges with respect to coexisting hosts has still not been definitively addressed.

Prevalence of specific and broad host range phages

Given the diversity of host genotypes, intuition suggests that a capacity to enter, and replicate in, diverse cell types would be advantageous in that it would increase the rate of successful encounters with victims. The need for the use of enrichments to detect phages for specific hosts was generally taken to indicate that marine phages are highly specific. Associations between specific host groups and phages are discussed in a later section but here several studies are highlighted for their large-scale perspective on the question of the prevalence of different breadths of host range.

A meta-analysis of 38 phage host range cross test matrices, representing ~12,000 possible infections, was conducted to assess whether there were common properties to

phage-host interaction networks (Flores et al., 2011). Most matrices were found to be nested rather than modular. Nestedness reflects phage host ranges that vary from “generalist” to “specialist”, host susceptibilities that range from highly susceptible to resistant, and a susceptibility of resistant hosts to generalist phages rather than specialist phages. Though matrices derived from diverse systems (only 9 of 38 from marine systems) this was shown not to have an impact on network structure and the overall median number of hosts per phage and phages per host were 6.13 and 3.04. The results of this meta-analysis suggest that broad host range phages are readily recovered. The presence of highly susceptible strains in these studies is consistent with the recovery and use of highly susceptible “indicator” strains by many workers. The fact that a single strain can be susceptible to many diverse viruses is more parsimoniously explained by a reduction in defensive systems than in an accumulation of diverse receptors, thus underscoring the likely importance of such defensive systems in contributing to host genotypic diversity in the environment. The lack of modularity in the majority of included studies is also suggestive of broad fundamental compatibility of viruses with host machinery necessary for their replication. One of the marine studies containing multiple representatives from 2 different genera within the same family (*Synechococcus* and *Prochlorococcus*) did identify multiple phages infecting both host groups (Sullivan et al., 2003). The authors point out that some modularity should reasonably be expected as a consequence of breakdowns in compatibility at some phylogenetic distance between potential hosts but that the limited scope of most of these studies with respect to the phylogenetic breadth of their host panels did not allow for detection of the bounds of such impact of phylogeny. Historically the existence of such modularity in phage infection patterns had however been suggested by observations made during the use of phages for typing of pathogens:

“In the case of a species which is already fairly well defined by several correlated characters, the testing of a number of phages on a group of strains of that species and of related species often shows the existence of ‘species-specific’ phages, that is phages which attack all (or nearly all) strains of one species, and no strains of any other species. The existence of

such phages is evidence of the value of the criteria used to define the species (Stocker 1955).”

With the primary aim of addressing whether geography played a role in structuring phage-host interactions Moebus and Nattkemper conducted host range assays between phages and hosts isolated primarily from the previously described Atlantic Ocean sampling, but also including some hosts from collections in the Bay of Biscay, and near Helgoland in the North Sea (Moebus and Nattkemper, 1981). These cross tests were between 774 bacterial strains and 298 bacteriophages, an unparalleled effort. Sequencing of these isolates was not a possibility at the time of these studies but Moebus identified clusters of strains related by phage-susceptibility patterns and later showed, as mentioned before, that these were in the *Vibrionaceae*. A high percentage of the host susceptibility profiles were unique (250/326 unique patterns just among the phage-sensitive Atlantic hosts), as were the infection profiles of the phages (224/258 unique infection patterns) (Moebus and Nattkemper, 1983).

Analysis of this matrix 30 years later yielded evidence for modularity at the large scale and both modularity and nestedness at the local scale (Flores et al., 2013). Analysis of the large scale modules indicated that although cross-infection between hosts and phages from different sampling sites was common, that interactions were three times more likely between members from the same site, suggesting some endemism in infection patterns. Altogether, the infections analyzed included 215 phages and 286 hosts, yielding 1332/61490 positive interactions. The mean number of phages infecting each host was 4.6 (max 20), the mean number of hosts infected by a phage was 6.19 (max 31), 43 (20%) phages infected only a single host in the panel and 41 (19%) infected over 10 hosts. As indicated previously, the majority of the host strains in this cross test were shown, on the basis morphological and metabolic characterization, to be in the family *Vibrionaceae*, with the majority in the genus *Vibrio* (280/326 phage sensitive strains examined) (Moebus and Nattkemper, 1983). Though Flores et al. had predicted that modular structure might be detectable in larger scale studies the modules detected here

can unfortunately not be evaluated in light of phylogeny due to the absence of sequence data for any of these strains.

Viral Differentiation in Life History Strategy

Host range is only one component of life history strategy, which also includes burst size, particle stability in the environment, adsorption rate, and latent period (De Paepe and Taddei, 2006). Assuming that the presence of a lineage in a given environment is evidence of a successful system of trade-offs then it appears that, despite apparently rare hosts, high specificity is a successful life history strategy for marine phages. It is then interesting to ask whether particular host range profiles are the purview of certain groups of phages. It has often been observed that with respect to host range often myoviridae are broad, siphoviridae are intermediate, and podoviridae are narrow (Sullivan et al., 2003; Wichels et al., 1998). Why there should be a relationship between structural morphology, genome content (beyond the structural elements), and host range is not clear. Though there is some mixing of replication modules across structural modules, it also appears generally that phage genomes are cohesive with morphotypes (Rohwer and Edwards, 2002). The fact that viruses of all morphotypes can generally be isolated for any host, and thus potentially co-infect, suggests that selective forces act against extensive genome mixing across phage morphotypes. It is possible that viral morphology is related to optimization for a particular ecological niche defined by the integration of physical conditions and host availability in relation to space and time.

***Vibrio* as a Model System**

The prevailing view of viral production is that it is associated with the most rapidly growing strains within a host group. Succession of microbial taxa during phytoplankton blooms, each reaching up to 20% of total abundance from backgrounds of <5%, suggests that dynamic cycles in growth rates and group abundance are a common feature for many groups of marine bacteria (Teeling et al., 2012). One of the most often cited of these low-abundance but quickly-responding groups, are the *Vibrio* (Giovannoni

et al., 2005; Suttle, 2007). Though they are generally present at low cell densities they are described as opportunitrophs for their characteristic ability to quickly respond to increased nutrient conditions (Eilers et al., 2000). The number of rRNA operons has been shown to correlate with capacity for rapid response to increased nutrient concentrations and whereas the average number of 16S in the rrnDB is 4.2 copies, the average copy number⁴ within the *Vibrionaceae* is ~10 (Klappenbach et al., 2000; Lee et al., 2009). Indeed, conditions that disproportionately favor *Vibrio* growth, such as nutrient amendments and higher water temperatures, are also associated with greater viral productivity in marine microbial communities. Even in the absence of bloom conditions several features of their cell physiology and behavior make *Vibrio* likely targets of phage infection. Some *Vibrio* are noted for their ability to become quite large - up to at least ~3.1µm in length and ~0.9µm in diameter, this provides for increasing access to diffusion-limited nutrients (Ovreas et al., 2003) but likewise increases encounter rate with viruses. Large cell size is also associated with larger viral burst sizes and these have been estimated to range up to at least 350 per cell for *Vibrio* under culture conditions (Hidaka and Tokushige, 1978). Cell motility acts as a mechanism for increasing encounter rates (Murray and Jackson, 1992), and whereas some of the most numerically abundant cells in the global ocean are nonmotile, such as *Prochlorococcus* and SAR11, many of the *Vibrio* are fast swimmers (Stocker, 2012) and thus exposed to larger encounter volumes. The addition of nutrients has been shown to increase the prevalence of motility in communities from background levels of ~10% to up to 80% (Stocker, 2012), consistent with the flourishing of *Vibrio* and underscoring that they are representative of the major shift in communities under such conditions.

In this work we focus on the *Vibrio* as representatives of the putative major contributors to viral productivity in marine microbial communities. A previously established model system for the *Vibrio* provides a framework within which novel environmental isolates can be readily assigned to previously identified populations of closely related strains with differentiated ecologies. Populations defined in this way

⁴ Copy numbers from rrnDB website (<http://rrnodb.umms.med.umich.edu/index.php>) accessed 04.06.14(14 *Vibrionaceae* of 1297 Bacteria in database with count of 16S copy number).

represent natural boundaries of gene flow and ecology. High rates of recombination within populations suggest low linkage among house keeping genes such that a phage compatible with components of the basic replication machinery in one host is likely compatible with the necessary proteins in another host within the population. Differences in rates of lateral gene transfer within and across populations define distinct regimes with respect to the frequency of mobile genes, such as phage receptors and defensive restriction-modification systems. And differences in host cell physiology or behavior associated with different ecologies represent different landscapes for optimization of phage life-history strategy (for example through trade-offs in burst size, latent period, robustness, host-range, or capacity for lysogeny).

Incorporation of population structure into the study of phage-host interactions in the environment thus affords the opportunity to assess the relationship between phage infection profiles on individual strains and the natural unit of bacterial evolution in the environment. This provides the basis for assessing whether the differentiation of hosts into distinct ecologies is also associated with distinct phage infection regimes with respect to predation intensity, infection strategies, or phage types.

GOALS OF THIS THESIS

The aim of this thesis is to provide quantitative estimates of fundamental aspects of virus-host interactions in the marine environment with respect to host population structure and ecology, and in light of viral genomes. To this end, methods are optimized to allow for increased throughput in handling phage in culture and for the concentration of infective viruses from seawater (Chapter 2, Appendix 1). A new large scale *Vibrio* phage-host system is developed and comprised of co-isolated genome-sequenced phages and hosts embedded in the context of a 93 consecutive day environmental time series (Chapter 3). Fundamental aspects of host-phage infection networks within a community are characterized with estimates of strain-level susceptibility to lytic phage predation and large-scale host range cross tests to determine prevalence of different phage infection profiles, and these results are considered in light of host population structure and phage genomes (Chapter 4).

REFERENCES

- Acinas, S.G., Klepac-Ceraj, V., Hunt, D.E., Pharino, C., Ceraj, I., Distel, D.L., Polz, M.F., 2004. Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 430, 551-554.
- Ahrens, R., 1968. Taxonomische Untersuchungen an sternbildenden *Agrobacterium*-Arten aus der westlichen Ostsee. *Kiel Meeresforsch* 24, 147-173.
- Ahrens, R., 1971. Untersuchungen zur Verbreitung von Phagen der Gattung *Agrobacterium* in der Ostsee. *Kieler Meeresforsch* 27, 102-112.
- Bergh, Ø., Børsheim, K.Y., Bratbak, G., Heldal, M., 1989. High abundance of viruses found in aquatic environments. *Nature* 340, 467-468.
- Biers, E.J., Wang, K., Pennington, C., Belas, R., Chen, F., Moran, M.A., 2008. Occurrence and expression of gene transfer agent genes in marine bacterioplankton. *Applied and environmental microbiology* 74, 2933-2939.
- Biller, S.J., Schubotz, F., Roggensack, S.E., Thompson, A.W., Summons, R.E., Chisholm, S.W., 2014. Bacterial vesicles in marine ecosystems. *Science* 343, 183-186.
- Bratbak, G., Heldal, M., Norland, S., Thingstad, T.F., 1990. Viruses as partners in spring bloom microbial trophodynamics. *Applied and environmental microbiology* 56, 1400-1405.
- Cho, B.C., Azam, F., 1988. Major role of bacteria in biogeochemical fluxes in the ocean's interior. *Nature* 332, 441-443.
- Chow, C.-E.T., Kim, D.Y., Sachdeva, R., Caron, D.A., Fuhrman, J.A., 2013. Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *The ISME Journal*.
- Claverie, J.-M., Abergel, C., 2010. Mimivirus: the emerging paradox of quasi-autonomous viruses. *Trends in Genetics* 26, 431-437.
- Cordero, O.X., Polz, M.F., 2014. Explaining microbial genomic diversity in light of evolutionary ecology. *Nature reviews. Microbiology* 12, 263-273.
- De Paepe, M., Taddei, F., 2006. Viruses' life history: towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS biology* 4, e193-e193.
- Eilers, H., Pernthaler, J., Amann, R., 2000. Succession of pelagic marine bacteria during enrichment: a close look at cultivation-induced shifts. *Applied and environmental microbiology* 66, 4634-4640.
- Espejo, R.T., Canelo, E.S., 1968. Properties of bacteriophage PM2: a lipid-containing bacterial virus. *Virology* 34, 738-747.
- Flores, C.O., Meyer, J.R., Valverde, S., Farr, L., Weitz, J.S., 2011. Statistical structure of host-phage interactions. *Proceedings of the National Academy of Sciences* 108, E288-E297.
- Flores, C.O., Valverde, S., Weitz, J.S., 2013. Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. *The ISME Journal* 7, 520-532.
- Forterre, P., Soler, N., Krupovic, M., Marguet, E., Ackermann, H.-W., 2012. Fake virus particles generated by fluorescence microscopy. *Trends in microbiology*.
- Fredericq, P., 1950. A Rapid Simple Technique for the Growth and Preservation of Enterobacteriophages. *Compte rendu des seances de la Societe de biologie* 144, 295-297.

Gilbert, J.A., Mühling, M., Joint, I., 2008. A rare SAR11 fosmid clone confirming genetic variability in the 'Candidatus Pelagibacter ubique' genome. *The ISME Journal* 2, 790-793.

Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., Bibbs, L., Eads, J., Richardson, T.H., Noordewier, M., 2005. Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309, 1242-1245.

Hennes, K.P., Suttle, C.A., Chan, A.M., 1995. Fluorescently Labeled Virus Probes Show that Natural Virus Populations Can Control the Structure of Marine Microbial Communities. *Applied and environmental microbiology* 61, 3623-3627.

Hidaka, T., 1977. Detection and isolation of marine bacteriophage systems in the southwestern part of the Pacific Ocean. *Mem. Fac. Fish. Kagoshima Univ* 26, 55-62.

Hidaka, T., Tokushige, A., 1978. Isolation and characterization of *Vibrio parahaemolyticus* bacteriophages in sea water. *Mem. Fac. Fish. Kagoshima Univ* 27, 79-90.

Hobbie, J.E., Daley, R.J., Jasper, S., 1977. Use of nuclepore filters for counting bacteria by fluorescence microscopy. *Applied and environmental microbiology* 33, 1225-1228.

Holmfeldt, K., Odic, D., Sullivan, M.B., Middelboe, M., Riemann, L., 2012. Cultivated single-stranded DNA phages that infect marine Bacteroidetes prove difficult to detect with DNA-binding stains. *Appl Environ Microbiol* 78, 892-894.

Hughes, S.S., 1977. *The virus: A history of the concept*. Heinemann educational books London.

Jannasch, H.W., Jones, G.E., 1959. Bacterial populations in sea water as determined by different methods of enumeration. *Limnol. Oceanogr* 4, 128-139.

Jost, G., Wiese, J., 2013. Temporal variations in the concentrations of bacteria and their lytic phages: an example of an indigenous phage host system in Lake Plusharpsee, Germany. *Fundamental and Applied Limnology/Archiv für Hydrobiologie* 182, 183-190.

Kausche, G.A., Pfankuch, E., Ruska, H., 1939. Die Sichtbarmachung von pflanzlichem Virus im Übermikroskop. *Naturwissenschaften* 27, 292-299.

Klappenbach, J.A., Dunbar, J.M., Schmidt, T.M., 2000. rRNA operon copy number reflects ecological strategies of bacteria. *Applied and environmental microbiology* 66, 1328-1333.

Lee, Z.M.-P., Bussema, C., Schmidt, T.M., 2009. rrnDB: documenting the number of rRNA and tRNA genes in bacteria and archaea. *Nucleic acids research* 37, D489-D493.

Long, R.A., Azam, F., 2001. Microscale patchiness of bacterioplankton assemblage richness in seawater. *Aquatic Microbial Ecology* 26, 103.

McDaniel, L.D., Young, E., Delaney, J., Ruhnau, F., Ritchie, K.B., Paul, J.H., 2010. High frequency of horizontal gene transfer in the oceans. *Science* 330, 50-50.

Mendes, C., Santos, L., Cunha, Â., Gomez, N.C., Almeida, A., 2013. Proportion of prokaryotes enumerated as viruses by epifluorescence microscopy. *Annals of Microbiology*, 1-6.

Moebus, K., 1980. A method for the detection of bacteriophages from ocean water. *Helgoländer Meeresuntersuchungen* 34, 1-14.

Moebus, K., 1992. Further investigations on the concentration of marine bacteriophages in the water around Helgoland, with reference to the phage-host systems encountered. *Helgoländer Meeresuntersuchungen* 46, 275-292.

Moebus, K., Nattkemper, H., 1981. Bacteriophage sensitivity patterns among bacteria isolated from marine waters. *Helgoländer Meeresuntersuchungen* 34, 375-385.

Moebus, K., Nattkemper, H., 1983. Taxonomic investigations of bacteriophage sensitive bacteria isolated from marine waters. *Helgoländer Meeresuntersuchungen* 36, 357-373.

Murray, A.G., Jackson, G.A., 1992. Viral dynamics: a model of the effects of size, shape, motion and abundance of single-celled planktonic organisms and other particles. *Marine ecology progress series*. Oldendorf 89, 103-116.

Needham, D.M., Chow, C.-E.T., Cram, J.A., Sachdeva, R., Parada, A., Fuhrman, J.A., 2013. Short-term observations of marine bacterial and viral communities: patterns, connections and resilience. *The ISME journal* 7, 1274-1285.

Noble, R., Fuhrman, J., 1997. Virus decay and its causes in coastal waters. *Applied and environmental microbiology* 63, 77-83.

Ovreas, L., Bourne, D., Sandaa, R.A., Casamayor, E.O., Benlloch, S., Goddard, V., Smerdon, G., Heldal, M., Thingstad, T.F., 2003. Response of bacterial and viral communities to nutrient manipulations in seawater mesocosms. *Aquatic Microbial Ecology* 31, 109-121.

Philippe, N., Legendre, M., Doutre, G., Couté, Y., Poirot, O., Lescot, M., Arslan, D., Seltzer, V., Bertaux, L., Bruley, C., 2013. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* 341, 281-286.

Proctor, L., 1991. Roles of viral infection in organic particle flux. *Mar. Ecol. Prog. Ser.* 69, 133-142.

Proctor, L.M., Fuhrman, J.A., 1990. Viral mortality of marine bacteria and cyanobacteria.

Rohwer, F., Edwards, R., 2002. The Phage Proteomic Tree: a genome-based taxonomy for phage. *Journal of Bacteriology* 184, 4529-4535.

Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., Wu, D., Eisen, J.A., Hoffman, J.M., Remington, K., 2007. The Sorcerer II global ocean sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS biology* 5, e77.

Sherr, E.B., 1989. Aquatic viruses. And now, small is plentiful. *Nature* 340, 429-429.

Spencer, R., 1955. A marine bacteriophage. *Nature* 175, 690-691.

Spencer, R., 1960. Indigenous marine bacteriophages. *Journal of Bacteriology* 79, 614.

Stocker, R., 2012. Marine Microbes See a Sea of Gradients. *Science* 338, 628-633.

Stocker, R., Seymour, J.R., Samandani, A., Hunt, D.E., Polz, M.F., 2008. Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches. *Proc Natl Acad Sci U S A* 105, 4209-4214.

Sullivan, M.B., Waterbury, J.B., Chisholm, S.W., 2003. Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. *Nature* 424, 1047-1051.

Suttle, C.A., 2007. Marine viruses--major players in the global ecosystem. *Nature Reviews Microbiology* 5, 801-812.

Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M., Kassabgy, M., Huang, S., Mann, A.J., Waldmann, J., 2012. Substrate-controlled

succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336, 608-611.

Thingstad, T.F., Lignell, R., 1997. Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquatic Microbial Ecology* 13, 19-27.

Thompson, J.R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D.E., Benoit, J., Sarma-Rupavtarm, R., Distel, D.L., Polz, M.F., 2005. Genotypic diversity within a natural coastal bacterioplankton population. *Science* 307, 1311-1313.

Torrella, F., Morita, R.Y., 1979. Evidence by electron micrographs for a high incidence of bacteriophage particles in the waters of Yaquina Bay, Oregon: ecological and taxonomical implications. *Applied and environmental microbiology* 37, 774-778.

Uchino, Y., Hirata, A., Yokota, A., Sugiyama, J., 1998. Reclassification of marine *Agrobacterium* species: Proposals of *Stappia stellulata* gen. nov., comb. nov., *Stappia aggregata* sp. nov., nom. rev., *Ruegeria atlantica* gen. nov., comb. nov., *Ruegeria gelatinovora* comb. nov., *Ruegeria algalicola* comb. nov., and *Ahrensia kieliense* gen. nov., sp. nov., nom. rev. *Journal of General and Applied Microbiology* 44, 201-210.

Vage, S., Storesund, J.E., Thingstad, T.F., 2013. Adding a cost of resistance description extends the ability of virus-host model to explain observed patterns in structure and function of pelagic microbial communities. *Environ Microbiol* 15, 1842-1852.

Våge, S., Storesund, J.E., Thingstad, T.F., 2013a. Adding a cost of resistance description extends the ability of virus--host model to explain observed patterns in structure and function of pelagic microbial communities. *Environmental microbiology* 15, 1842-1852.

Våge, S., Storesund, J.E., Thingstad, T.F., 2013b. SAR11 viruses and defensive host strains. *Nature* 499, E3-E4.

Valentine, A.F., Chapman, G.B., 1966. Fine structure and host-virus relationship of a marine bacterium and its bacteriophage. *Journal of Bacteriology* 92, 1535.

Verdugo, P., Alldredge, A.L., Azam, F., Kirchman, D.L., Passow, U., Santschi, P.H., 2004. The oceanic gel phase: a bridge in the DOM-POM continuum. *Marine Chemistry* 92, 67-85.

Vergin, K.L., Tripp, H.J., Wilhelm, L.J., Denver, D.R., Rappé, M.S., Giovannoni, S.J., 2007. High intraspecific recombination rate in a native population of *Candidatus Pelagibacter ubique* (SAR11). *Environmental Microbiology* 9, 2430-2440.

Wang, K., Wommack, K.E., Chen, F., 2011. Abundance and distribution of *Synechococcus* spp. and cyanophages in the Chesapeake Bay. *Applied and environmental microbiology* 77, 7459-7468.

Waterbury, J.B., Valois, F.W., 1993. Resistance to co-occurring phages enables marine *synechococcus* communities to coexist with cyanophages abundant in seawater. *Appl Environ Microbiol* 59, 3393-3399.

Watson, S., Novitsky, T., Quinby, H.L., Valois, F., 1977. Determination of bacterial number and biomass in the marine environment. *Applied and environmental microbiology* 33, 940-946.

Wichels, A., Biel, S.S., Gelderblom, H.R., Brinkhoff, T., Muyzer, G., Schütt, C., 1998. Bacteriophage diversity in the North Sea. *Applied and environmental microbiology* 64, 4128-4133.

- Wilhelm, L.J., Tripp, H.J., Givan, S.A., Smith, D.P., Giovannoni, S.J., 2007. Natural variation in SAR11 marine bacterioplankton genomes inferred from metagenomic data. *Biol Direct* 2, 27.
- Wilhelm, S.W., Weinbauer, M.G., Suttle, C.A., Jeffrey, W.H., 1998. The role of sunlight in the removal and repair of viruses in the sea. *Limnology and Oceanography* 43, 586-592.
- Wommack, K.E., Colwell, R.R., 2000. Virioplankton: viruses in aquatic ecosystems. *Microbiology and molecular biology reviews* : MMBR 64, 69-114.
- Wommack, K.E., Hill, R.T., Muller, T.A., Colwell, R.R., 1996. Effects of sunlight on bacteriophage viability and structure. *Applied and environmental microbiology* 62, 1336-1341.
- Wommack, K.E., Ravel, J., Hill, R.T., Chun, J., Colwell, R.R., 1999. Population dynamics of Chesapeake Bay virioplankton: total-community analysis by pulsed-field gel electrophoresis. *Applied and environmental microbiology* 65, 231-240.
- Yaron, S., Kolling, G.L., Simon, L., Matthews, K.R., 2000. Vesicle-Mediated Transfer of Virulence Genes from *Escherichia coli* O157: H7 to Other Enteric Bacteria. *Applied and environmental microbiology* 66, 4414-4420.
- Zachary, A., 1974. Isolation of bacteriophages of the marine bacterium *Beneckea natriegens* from coastal salt marshes. *Applied microbiology* 27, 980-982.
- Zhao, Y., Temperton, B., Thrash, J.C., Schwalbach, M.S., Vergin, K.L., Landry, Z.C., Ellisman, M., Deerinck, T., Sullivan, M.B., Giovannoni, S.J., 2013. Abundant SAR11 viruses in the ocean. *Nature* 494, 357-360.

Chapter Two

Streamlining standard bacteriophage methods for higher throughput

ABSTRACT

The application of culture based methods to the quantitative study of environmental phage-host systems is limited because of the generally low abundance of specific phages and the high diversity of interacting members. As a result, studies of environmental phage-host interactions are often confounded by the need to use methods such as enrichments or previously isolated hosts. Broadly accessible methods for recovery of highly concentrated infective viruses, and their characterization in culture, would facilitate more widespread use of large-scale and environmentally representative approaches. Here, we present a series of optimized and novel methods for increasing efficiency and throughput in the isolation and culture-based study of environmental phage-host systems. The agar overlay, the universal tool of the phage worker, is streamlined to eliminate steps and materials usage. Likewise, serial purification of phages is optimized to eliminate unnecessary steps by adapting methods used for serial colony purification of bacteria. A novel method is presented for recovery of highly concentrated infective viruses using oxalate resuspension of iron precipitated seawater. A general approach for recovery of environmental phage-host systems using these streamlined and novel methods is presented that is based on direct plating of iron-oxalate concentrated viruses in agar overlays. Recommendations are made for efficient archival and stock development of newly isolated phages. And finally, an optimized spot-test for high throughput determination of host range is presented. It is hoped that in aggregate these methods facilitate an increase in scale for workers interested in representative recovery of marine phage-host systems.

INTRODUCTION

Phages exert their influence on microbial systems through infections of host cells. Though diverse tools have been brought to bear in the study of viruses, culture based approaches are still the most powerful way to determine whether specific phages are capable of infecting and lysing specific hosts. However, several aspects of phage host interactions in the environment pose a challenge to culture-based studies in this context. Firstly, phages for any given host are generally rare, and secondly, both phages and hosts are highly diverse such that studies with small numbers of hosts may yield limited success in terms of recovery of specific phages. For these reasons, methods such as enrichments or assays with highly sensitive allochthonous host

strains (“indicator strains”) are often used. These methods do not yield quantitative representations of the phage-host interactions that are occurring *in situ*. The availability of simple and low cost methods for recovery of highly concentrated infective environmental viruses, and streamlining of methods for dealing with these phage-host systems in a higher throughput fashion in culture, would facilitate a more widespread use of environmentally representative culture-based studies of phage-host infections.

Here, we provide a guide intended to facilitate large scale quantitative isolation and characterization of novel environmental phage-host systems. We recommend methods, and for some of these methods we present new modifications and optimizations that increase efficiency and throughput. Use of these methods allows for quantitative characterization of the number of infective viruses present per host strain as well as subsequent isolation and host-range characterization of these viruses. When used with viruses and hosts co-isolated from the same sample they provide quantitative representations of *in situ* host-phage interaction networks. We first discuss the most commonly used phage methods and present optimizations that improve their efficiency. We then highlight a series of phage methods that are particularly relevant for application to development of novel environmental phage-host model systems on a large scale. Throughout, we refer to a *Vibrio* model system to provide concrete examples of application of the methods described.

RESULTS & DISCUSSION

Section I. Improvements in the agar overlay & use in rapid serial purification

The techniques for optimizing plaque formation on bacterial lawns were developed in the 1930s and 1940s. The most commonly cited reference for this method is its presentation in the book *Bacteriophages* by Mark Adams, in this book Adams indicates that “the agar layer for plating bacterial viruses seems to have been first described by Gratia (1936), and is in general use by practically all workers” (Adams, 1959; Gratia, 1936).¹ This basic method appears under

¹ The presentation of this method in *Bacteriophages* was in an appendix that was essentially a reprint (with minor modifications) of a paper of Adams’ from 1950. It is interesting to compare citations, as determined based on a Google Scholar search in April 2014, for these papers: Gratia ‘s 1936 paper indicated by Adams to be the first presentation of the method (21), Adams’ 1950 paper (587), Adams’ 1959 book (3890) [Adams, M.H., 1950. Methods of study of bacterial viruses. Methods in medical research 2, 1050, Adams, M.H., 1959. Bacteriophages.

various names but all refer to the two layers of agar that are associated with layering a top agar, containing mixed bacteria and phage, over a bottom agar. The thin top agar layer allows penetration of oxygen, and its lower percentage agar allows for growth of cells and diffusion of phage particles. The bottom agar provides a source of nutrients to cells in the top layer that sustains growth long enough for phage to complete several cycles of replication and form macroscopic plaques. The infection of a bacterial cell by a single lytic phage can be observed as a zone of clearing, a plaque, in the confluent lawn of host bacteria. Nearly 70 years later, the agar overlay holds fast as a mainstay in bacteriophage laboratories (Kropinski et al., 2009; Sambrook and Russell, 2006), and continues to provide a reliable method for determining titers, isolating novel phages, and purifying phage plaques. The quantitative relationship between bacteriophage particles and macroscopic plaques allows simple agar overlay plating techniques to provide consistent estimates of bacteriophage titer in stocks of unknown concentration. While these estimates are often lower than those made by flow cytometry or microscopy based methods, the plaque assay is the only method that can measure infectivity (Anderson et al., 2011). The continued publication of novel optimizations to improve efficiency, plaque detectability, and tailoring to specific host-phage systems, speak to the power and widespread use of this method (Hurst et al., 1994; Islam et al., 2012; Lillehaug, 1997; Pattee, 1966; Rizvi and Mora, 1963; Serwer et al., 2007). By implementing the slight modifications to standard practice described here, the efficiency of agar overlay plating, and serial purification, can be appreciably improved.

Streamlining the agar overlay. The current standard practice in preparation of agar overlay assays is as follows: bacteria and phage are added to a tube containing 3mL of molten agar, the contents are mixed, and the mixture is poured into a petri dish containing bottom agar. We have found that by simply pipetting the host culture, phage, and top agar directly onto the bottom agar it is possible to reduce the amount of materials and time required to plate large numbers of agar overlays (Table 1, Fig. 1).² The top agar, once added, is distributed by vigorous but brief swirling to ensure that it is distributed across the entirety of the plate. The benefits of this

Bacteriophages, Gratia, A., 1936. Numerical relations between lysogenic bacteria and particles of bacteriophage. Ann. Inst. Pasteur 57, 72.]

² Surprisingly, reference to a 1943 paper of Hershey et al. where they describe use of the agar overlay method reveals that they indicate that “A further slight improvement may be made by mixing the sample directly on the phage with only 3 ml 0.7 per cent agar, but the mixing is difficult.” This appears to be precisely the tube-free method approach presented here but it did not come strongly recommended and thus was apparently never adopted. (Hershey, A., Kalmanson, G., Bronfenbrenner, J., 1943. Quantitative methods in the study of the phage-antiphage reaction. J. Immunol 46, 267-279.)

approach include simplified experimental set-up and reductions in preparation times and material requirements. In addition, by eliminating tube-based steps, the duration of exposure to heat of phage and bacteria is reduced and the need for potentially damaging vortexing or agitation of phage mixtures is eliminated. Pairwise comparisons of the titers when plating vibriophages of three different morphologies (12G01 phage, myovirus; *Vibrio* phage 12G5 Jenny, siphovirus; *Vibrio* phage A1, podovirus) show no significant differences ($p > 0.01$) in titer between the two methods (Table 2). Using this streamlined, tubeless, plating method, over 45 samples can be plated per hour from a common or pre-prepared phage stock, without the need for individual tubes of molten agar or multiple transfers of bacteria and phage. When deployed for isolation of novel phages, or quantification of environmental viruses, this translates to a potential for screening hundreds of bacterial strains per day using a single bottle of molten agar rather than hundreds of tubes.

Serial purification. The tubeless approach to agar overlay plating can also be applied to speed serial purification of phage plaques. Purification of a phage is achieved, in standard practice, by performing three serial passages each derived from a single plaque. This is equivalent to triple-restreaking for colony purification of bacteria and by using the tubeless agar overlay method it is possible to perform the purification in an exactly analogous way by streaking from one plaque directly into a molten agar overlay. Various methods have been described by different workers but currently the most rapid are those that involve streaking of phage material from a plaque or liquid stock onto a bottom agar and then pouring on the mixture of molten top-agar and host bacteria and distributing with minimal agitation (PhageHunting_Program; Van Twest and Kropinski, 2009).

By streaking into a molten agar overlay that already contains the dispersed host it is possible to eliminate the use of tubes, achieve an even host lawn without spreading the pre-streaked phage, minimize handling of phage material, and reduce the time necessary for each serial purification. Streaking into molten top agar also leaves visible tracks in the agar that indicate where phages were applied, providing increased confidence that plaques observed derive from the applied sample rather than from autoinduction events wherein resident prophage form plaques. Using this approach, 45 phages can be comfortably re-streaked per hour and, in systems where plaques develop by the next day, these 45 phage can be triply purified in three days. Thus, in tubeless streaking-for-singles, serial purification of phage plaques can be achieved rapidly and

with minimal materials by streaking with a toothpick from a plaque-stab directly into a molten top agar already containing dispersed hosts (Table 3, Fig. 2).

The protocols described here, by virtue of their simplicity, reduce the time, steps, and materials necessary for both agar overlay plaque assays and serial purification of bacteriophage. These modifications to standard practice simplify experimental set-up, reduce preparation time, require fewer materials, and facilitate increased experimental scaling.

Section II. Isolation and characterization of environmental viruses

Isolation of novel viruses from the environment is commonly achieved using enrichment cultures. Enrichment cultures work well because they take advantage of the great diversity of viruses in natural samples and compensate for the low abundance of any one kind of virus by supplementation with large numbers of rapidly growing hosts. These rapidly growing hosts provide for sufficient encounter rates with even low abundance viruses and thus rapidly amplify their numbers. If the intent is simply to isolate a virus for a particular host then an enrichment may be satisfactory. However, enrichments cannot provide a quantitative representation of the types of viruses that are present in a sample and they add the risk of enriching for particular classes of viruses. Bias in enrichment can be a function of initial concentrations, such that more abundant viruses have a head-start in replication and thus achieve greater numbers. Bias may also be a function of life-history strategy and so, depending on the time given for enrichment to occur, may lead to selective enrichment of viruses with shorter latent periods or greater bursts sizes. Methods that can provide a less-biased representation of all viruses that are present are thus desirable and generally recommended. Direct plating of samples containing viruses into an agar overlay of hosts, as described above, is often used where quantitative estimates are desired but generally require the use of concentrated environmental samples to compensate for the low concentrations of specific viruses. In the following sections we describe the use of efficient methods for concentration of viruses and direct plating for quantitative isolation of naturally occurring phages. We close by presenting optimization for higher throughput of the commonly used spot test for host range.

Concentration of infective viruses by iron chloride flocculation. Iron chloride flocculation provides a method for achieving 1000-fold concentrations of viruses in marine water samples at low cost and using fairly standard laboratory equipment. Viruses are concentrated by pre-filtration of samples to remove bacterial cells and addition of an iron chloride solution. The formation of iron hydroxide polymers incorporates viral particles and these large flocculants can then be readily captured on 0.2µm filters commonly used to retain bacterial cells. Dissolution of these captured particulates liberates the viral particles.

The chemistry used to induce dissolution of the flocculants impacts viral infectivity. The optimization of the iron co-precipitation protocol for recover of infective viruses was conducted by this author, is included in (John et al., 2011), and is described in greater detail here. As originally formulated, an ascorbate solution is used as a reducing agent to dissolve iron flocculants. Dissolution in ascorbate is rapid and complete but results in loss of infectivity of viruses. Ascorbate solutions are also chemically unstable and different samples treated in the same way could be observed sometimes to change in color, ranging from dark purple to yellowish, over time or not at all. Treatment with ascorbate reduces iron(III) in hydroxides, dissolves them, and releases iron(II) which is then chelated by EDTA in solution. Iron(II) in complex with EDTA participates in Fenton reactions to produce hydroxyl radicals. Production of radical species has been shown to damage viruses and the combination of ascorbate with transition metals has been shown to be especially effective at catalyzing such reactions; the Fe-EDTA formed during the dissolution of iron precipitates in this method is itself “an excellent catalyst of ascorbate oxidation” (Buettner and Jurkiewicz, 1996; Turner, 1964). A variety of approaches were thus explored to minimize the impacts of such reactions on precipitated viruses resuspended in ascorbate, these included the following: addition of catalase to scavenge oxygen radicals, treatment under dark conditions to prevent photolytic oxidation of ascorbate, exchange into saline buffer using centrifugal dialysis. None of these methods were effective in preventing loss of infective viruses and exchanges of resuspended viral precipitates using saline solutions resulted in formation of new precipitates of unknown composition. Review of alternative chemistries indicated that oxalate was a potential candidate because of an apparently ligand-mediated, rather than reductive, mechanism for dissolution of iron hydroxide polymers. Both the use of ascorbate and oxalate solutions to dissolve iron had previously been used in the literature to clean phytoplankton surfaces of iron during studies of iron limitation, where strict assignment

of iron concentrations to intra- and extracellular pools was necessary (Tovar-Sanchez et al., 2003). Oxalate had also been tested in preliminary test in the development of the viral iron co-precipitation method but had been dismissed in favor of ascorbate due to slower dissolution rates. Following identification of oxalate as a potential candidate for preservation of viral infectivity it was tested by incorporating it into the resuspension solution in place of ascorbic acid. The use of oxalate yielded recovery of infective viruses that could be stored at 4°C over at least 38 days (the maximum test) with no significant decrease in titer (Figure 3). Tests with 2 *Vibrio* phages showed average initial recoveries of approximately 52%, however a test with a cyanomyovirus by Deng showed recovery of only 13% of infective viruses, suggesting that recoveries may differ across classes of virus [(John et al., 2011), Appendix 1].

Evaluation of the diversity within the Nahant Collection of vibriophages, which were isolated and brought into culture using iron co-precipitation followed by oxalate resuspension, suggests that they are highly diverse. All morphotypes of the *Caudovirales*, the *Myoviridae*, *Podoviridae*, and *Siphoviridae*, are represented in the collection, as assessed by either electron microscopy or genome sequence homology to other viruses of the morphotype; as are nontailed-like viruses, identified as such on the basis of sequence homology (detailed methods on phage isolation and preparation for sequencing, and genome analysis are described in Chapter 3). Additional viruses isolated as part of this isolation effort presented difficulties in preparation for genome sequencing and thus were not sequenced, these may represent non-dsDNA viruses and thus additional diversity, though this remains to be tested. To assess the relative diversity of Fe-oxalate recovered viruses with respect to that of known viruses a well-represented gene was used to construct a combined phylogeny with publically available virus sequences. The terminase large subunit (LSU) was present in 264 viruses in the collection and was compared with 830 terminase LSU proteins available in the ACLAMEdb of mobile genetic elements (Figure 4) and were shown to be representative of the diversity present in this large public collection as well as including additional clades unique to the Nahant Collection (Leplae et al., 2004; Leplae et al., 2010). Sequences were aligned using muscle, FastTree was used to infer an approximate-maximum likelihood phylogenetic tree, the tree was visualized using iTOL (Edgar, 2004a, b; Letunic and Bork, 2007, 2011; Price et al., 2009, 2010).

It is important to acknowledge the loss of large and particulate-associated viruses that are systematically removed with 0.2µm pre-filtration, this is a very common feature of methods that seek to separate viruses from cells and introduces an important operational bias (the reader is referred to Chapter 1 for additional discussion of operational choices in sampling for viruses). Although separation of cells subsequent to flocculation and dissolution is possible this introduces the risk reducing infective viruses by increasing their encounter rates with highly concentrated cells. Future studies to compare differential compositions of viruses isolated from the pre-filtered fraction with those collected in the filtrate would be useful to assess whether there are systematic biases.

Direct plating for quantitatively representative isolation. Viral concentrates collected by iron flocculation can be used directly in agar-overlay assays. Using concentrates in such agar overlays yields distinct plaques for each infective virus present and thus provides both a measure of predator load for each strain and a source from which to cultivate novel viral isolates. For example, given 1000-fold concentration it is possible to estimate the number of viral predators present in 15mL of source water by simply incorporating 15µL of concentrate into an agar overlay. It is important, in all such assays to include two types of negative controls, the first containing no concentrate and the second containing a virus-free concentrate prepared in the same way. These controls ensure that auto-induction, the formation of plaques by induced resident prophage in the absence of added virus, in response to plating conditions or the addition of the iron-oxalate resuspension solution can be detected. Diversity of recovered viruses is increased by holding the agar overlay platings of viral concentrates under observation for extended periods of time. In our experience plates maintained so as to prevent desiccation can be seen to yield new plaques even after 30 days and may represent groups of viruses that are commonly missed using standard hold times of only a day or several days.

Archiving and storage of large numbers of novel viral isolates. Whereas in some small-scale studies it may be sufficient to proceed directly from direct plating of concentrates to serial purification of novel isolates, as described in previous sections, in larger scale studies it may be desirable to store larger numbers of plaques for later study. Using the following approach it is possible to archive a large collection of plaques using standardized procedures (Table 4). Plaques are marked and numbered on each plate as they appear during the course of the agar overlay

incubation and once the incubation period has ended all plaques are recorded in a master catalog that includes observations of plaque morphology (for example, Figure 5). Plaques are then picked, eluted, and processed using 96-well plate format processing. If a 96-well format filtration apparatus is available then it is possible to generate 0.2 μ m-filtered cell-free plaque eluates for storage at 4°C by filtering an aliquot of plaque eluate, otherwise plaque eluate is simply mixed with 50% glycerol and stored at -20°C. These steps can easily be applied to smaller numbers of samples by using single microcentrifuge tubes rather than 96-well plates. It is cautioned that phages differ in their robustness to different forms of storage and it is always recommended to store replicates of phage stocks at multiple conditions. Options for storage include 4°C for raw lysates, -20°C and -80°C for lysates mixed with glycerol, and lyophilization. An additional novel approach, based on storage of viruses as infected cells, was recently described and consists of exposing host cells to phages and then storing these infected cells as -80°C glycerol stocks (Golec et al., 2011). The reader is referred to Fortier et al. 2009 for a thorough review containing a variety of recommended protocols (Fortier and Moineau, 2009).

Preparation of high titer stocks of viruses. Preparation of high titer stocks requires first that novel phages have been serially purified using the methods described in previous sections, this may be done either by proceeding directly from the isolation plate plaques, or by using archived plaques stored as described in the preceding section. Both plate lysates and liquid lysates are commonly used methods for the production of high titer stocks for known phage-host systems (Swanstrom and Adams, 1951). Although liquid lysates are convenient when working with phages where growth conditions are known, they present a challenge when working with novel phages. Lytic interactions identified by plaque formation often fail to present clearly when observed in liquid culture. This may be due to dynamic cycles of resistant host cell lineages that mask lysis. If these cultures are not carefully tracked over time it is not clear whether they have not yet lysed or whether they have lysed and regrown. It is also not straight forward to determine the appropriate concentration of phage to add as different phage host systems may differ significantly in their burst size or latent period and thus require very different initial concentrations to achieve high titer lysates. For these reasons, we recommend the use of small scale enrichments of purified phages followed by large-size petri plate agar overlay lysates, as outlined in Table 6. This approach, combining small scale primary enrichments with a preliminary screening to determine appropriate titer for larger plate lysates, increases the success

rate when working with large numbers of unknowns and minimizes the need for repeated attempts to optimize conditions in liquid lysates. Plaques generated by serial purification are eluted overnight before being spiked with a small volume of actively growing host culture and incubated to generate a small scale primary lysate. Dilution series of primary lysates can then be prepared and spotted onto host agar overlay lawns to determine appropriate dilutions of primary lysate to use to generate nearly-confluent lysis of agar overlay lawns for harvest as high titer phage stocks.

Host range assays by high throughput spot test. Determination of viral host-range can be accomplished by incorporation of diluted lysates into agar overlays to allow for the formation of discrete plaques dispersed throughout the host lawn. Although incorporation into agar overlay remains the preferred method when working with small numbers of phage-host pairs it is a limiting step when scaling up. An often employed alternative method is the “spot-test”, the spot-test consists of spotting a small volume of lysate directly onto an agar overlay of the host lawn and observing for lysis. Criticisms of this method include the potential for killing-from-without, wherein viral titers are so high that cells are lysed without the replication viruses and thus introduce the risk of false positives. However, the use of the spot test affords significant scalability and the risk of false positives can be reduced by requiring that observed clearings show increases in size with time. In Table 7 we outline a method for scaling up the spot-test using 96-well blotters applied to host lawn agar overlays in oversized petri dishes. By using microwave steam sterilization of blotters we allow for rapid cycling of equipment necessary for plating such that 3 replicates each of 31 viruses, and a negative control, can be spotted onto 100 host lawns in a single plating session (3,200 replicated crosses). Considering that a recent meta-study of 38 host range matrices included a total of ~12,000 possible interactions, the method described here clearly provides a simple approach for substantially increasing throughput of such assays.

CONCLUSION

In aggregate, the methods described here provide an approach for efficient and large-scale quantitative studies of marine phage-host systems. Enumeration of the number of phage predators per host strain in the environment is not common in the literature but using the described methods in conjunction with concurrent host isolation makes this readily achievable. It

is hoped that the improvements in efficiency and general approach described will facilitate increased scaling by future workers describing novel environmental phage-host systems.

ACKNOWLEDGEMENTS

We thank Radhey Shyam Sharma and Simon Labrie for comments, suggestions, and valuable discussion on early versions of this manuscript.

REFERENCES

- Adams, M.H., 1950. Methods of study of bacterial viruses. *Methods in medical research* 2, 1050.
- Adams, M.H., 1959. Bacteriophages. *Bacteriophages*.
- Anderson, B., Rashid, M.H., Carter, C., Pasternack, G., Rajanna, C., Revazishvili, T., Dean, T., Senecal, A., Sulakvelidze, A., 2011. Enumeration of bacteriophage particles: Comparative analysis of the traditional plaque assay and real-time QPCR-and NanoSight-based assays. *Bacteriophage* 1, 86-93.
- Buettner, G.R., Jurkiewicz, B.A., 1996. Catalytic metals, ascorbate and free radicals: combinations to avoid. *Radiation research* 145, 532-541.
- Edgar, R.C., 2004a. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113.
- Edgar, R.C., 2004b. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 32, 1792-1797.
- Fortier, L.-C., Moineau, S., 2009. Phage production and maintenance of stocks, including expected stock lifetimes, *Bacteriophages*. Springer, pp. 203-219.
- Golec, P., Dąbrowski, K., Hejnowicz, M.S., Gozdek, A., Łoś, J.M., Węgrzyn, G., Łobocka, M.B., Łoś, M., 2011. A reliable method for storage of tailed phages. *Journal of microbiological methods* 84, 486-489.
- Gratia, A., 1936. Numerical relations between lysogenic bacteria and particles of bacteriophage. *Ann. Inst. Pasteur* 57, 72.
- Hershey, A., Kalmanson, G., Bronfenbrenner, J., 1943. Quantitative methods in the study of the phage-antiphage reaction. *J. Immunol* 46, 267-279.
- Hurst, C.J., Blannon, J.C., Hardaway, R.L., Jackson, W.C., 1994. Differential effect of tetrazolium dyes upon bacteriophage plaque assay titers. *Applied and environmental microbiology* 60, 3462-3465.
- Islam, M.R., Ogura, Y., Asadulghani, M., Ooka, T., Murase, K., Gotoh, Y., Hayashi, T., 2012. A sensitive and simple plaque formation method for the Stx2 phage of *Escherichia coli* O157: H7, which does not form plaques in the standard plating procedure. *Plasmid* 67, 227-235.
- John, S.G., Mendez, C.B., Deng, L., Poulos, B., Kauffman, A.K., Kern, S., Brum, J., Polz, M.F., Boyle, E.A., Sullivan, M.B., 2011. A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environ Microbiol Rep* 3, 195-202.
- Kropinski, A.M., Mazzocco, A., Waddell, T.E., Lingohr, E., Johnson, R.P., 2009. Enumeration of bacteriophages by double agar overlay plaque assay. *Methods Mol Biol* 501, 69-76.
- Leplae, R., Hebrant, A., Wodak, S.J., Toussaint, A., 2004. ACLAME: a CLAssification of Mobile genetic Elements. *Nucleic acids research* 32, D45-D49.
- Leplae, R., Lima-Mendez, G., Toussaint, A., 2010. ACLAME: a CLAssification of Mobile genetic Elements, update 2010. *Nucleic acids research* 38, D57-D61.
- Letunic, I., Bork, P., 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127-128.
- Letunic, I., Bork, P., 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic acids research* 39, W475-W478.
- Lillehaug, D., 1997. An improved plaque assay for poor plaque - producing temperate lactococcal bacteriophages. *Journal of applied microbiology* 83, 85-90.
- Pattee, P., 1966. Use of tetrazolium for improved resolution of bacteriophage plaques. *Journal of Bacteriology* 92, 787.
- PhageHunting_Program, Plaque Purification. University of Pittsburgh Phagehunting Program

Price, M.N., Dehal, P.S., Arkin, A.P., 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular biology and evolution* 26, 1641-1650.

Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5, e9490.

Rizvi, S., Mora, P.T., 1963. Bacteriophage plaque-count assay and confluent lysis on plates without bottom agar layer.

Sambrook, J., Russell, D.W., 2006. Plating Bacteriophage λ . *Cold Spring Harbor Protocols* 2006, pdb.prot3961.

Serwer, P., Hayes, S.J., Thomas, J.A., Hardies, S.C., 2007. Propagating the missing bacteriophages: a large bacteriophage in a new class. *Virology* 4, 21.

Swanstrom, M., Adams, M., 1951. Agar layer method for production of high titer phage stocks. *Royal Society of Medicine*, pp. 372-375.

Tovar-Sanchez, A., Sañudo-Wilhelmy, S.A., Garcia-Vargas, M., Weaver, R.S., Popels, L.C., Hutchins, D.A., 2003. A trace metal clean reagent to remove surface-bound iron from marine phytoplankton. *Marine Chemistry* 82, 91-99.

Turner, G., 1964. Inactivation of vaccinia virus by ascorbic acid. *Journal of General Microbiology* 35, 75-80.

Van Twest, R., Kropinski, A.M., 2009. Bacteriophage enrichment from water and soil, *Bacteriophages*. Springer, pp. 15-21.

FIGURES AND TABLES

Figure 1. Steps in tubeless agar overlay plating. Dispensing bacteria & phage onto bottom agar (A), molten agar set up (B), dispensing top agar directly onto bottom agar (C), and evenly spread plaques next day (D).

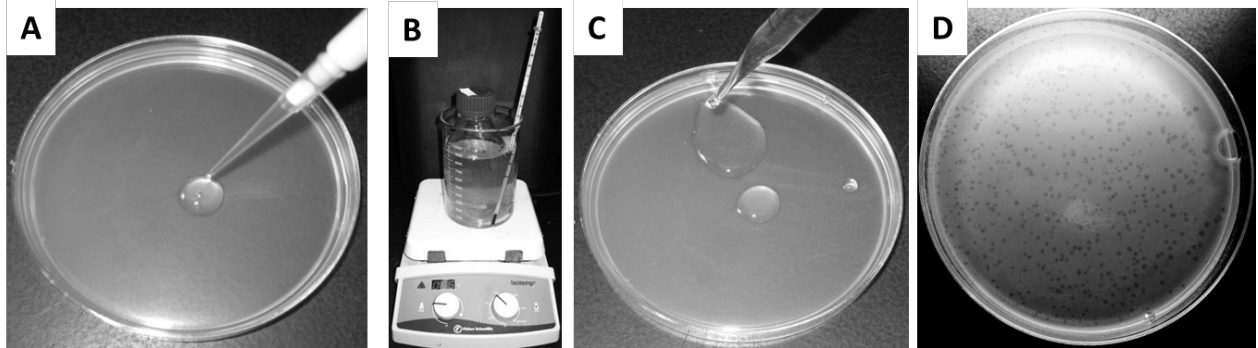




Figure 2. Image of single plaques generated by streaking-for-singles into molten top agar.

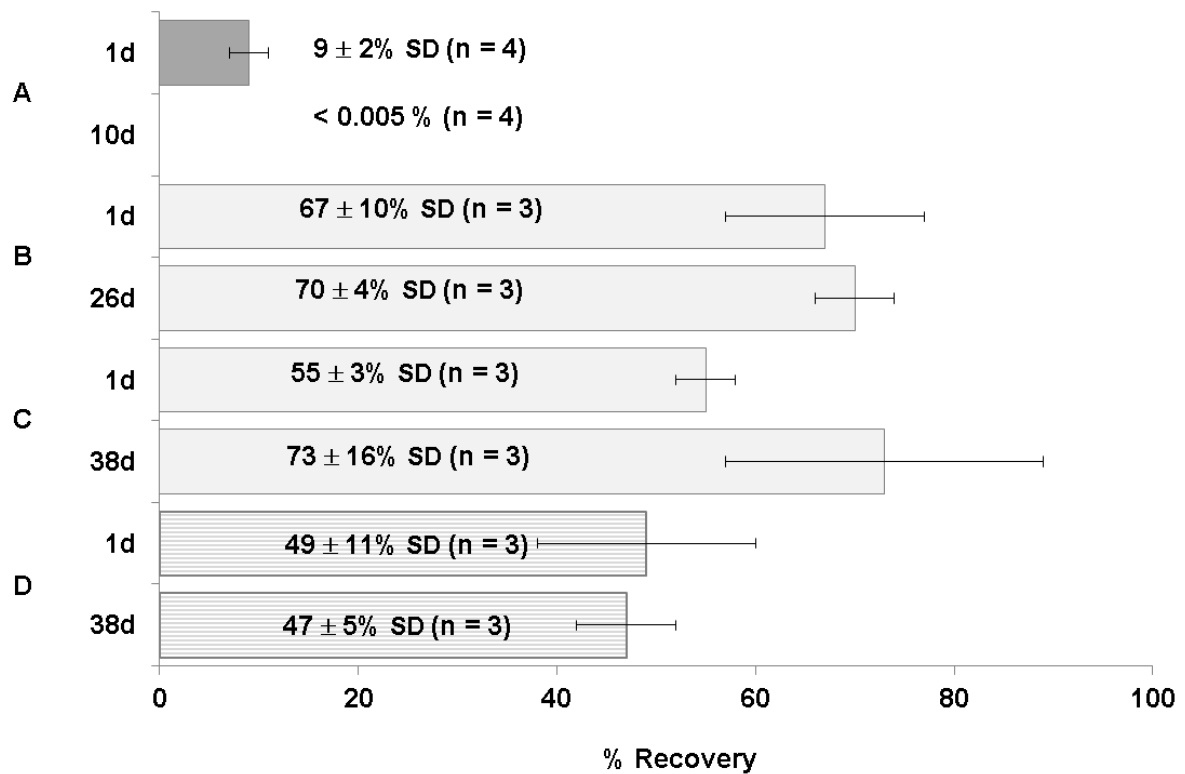


Figure 3. Recovery of infective virus in plaque forming units over time, following iron flocculation and resuspension in either ascorbate or oxalate solution. (A) & (B) in artificial seawater, (C) & (D) in 0.2uM-filtered aged natural seawater (negative for the presence of infective viruses for the target host). Different resuspension solutions indicated by shading: (A) resuspended in ascorbate [dark shading], (B, C, D) in oxalate solutions [light shading]. Virus morphology indicated by fill: (A, B, C) myovirus *Vibrio* phage 12G01 [solid] (D) siphovirus *Vibrio* phage 12G5 [striped]. As presented in Appendix 1 (John et al., 2011)

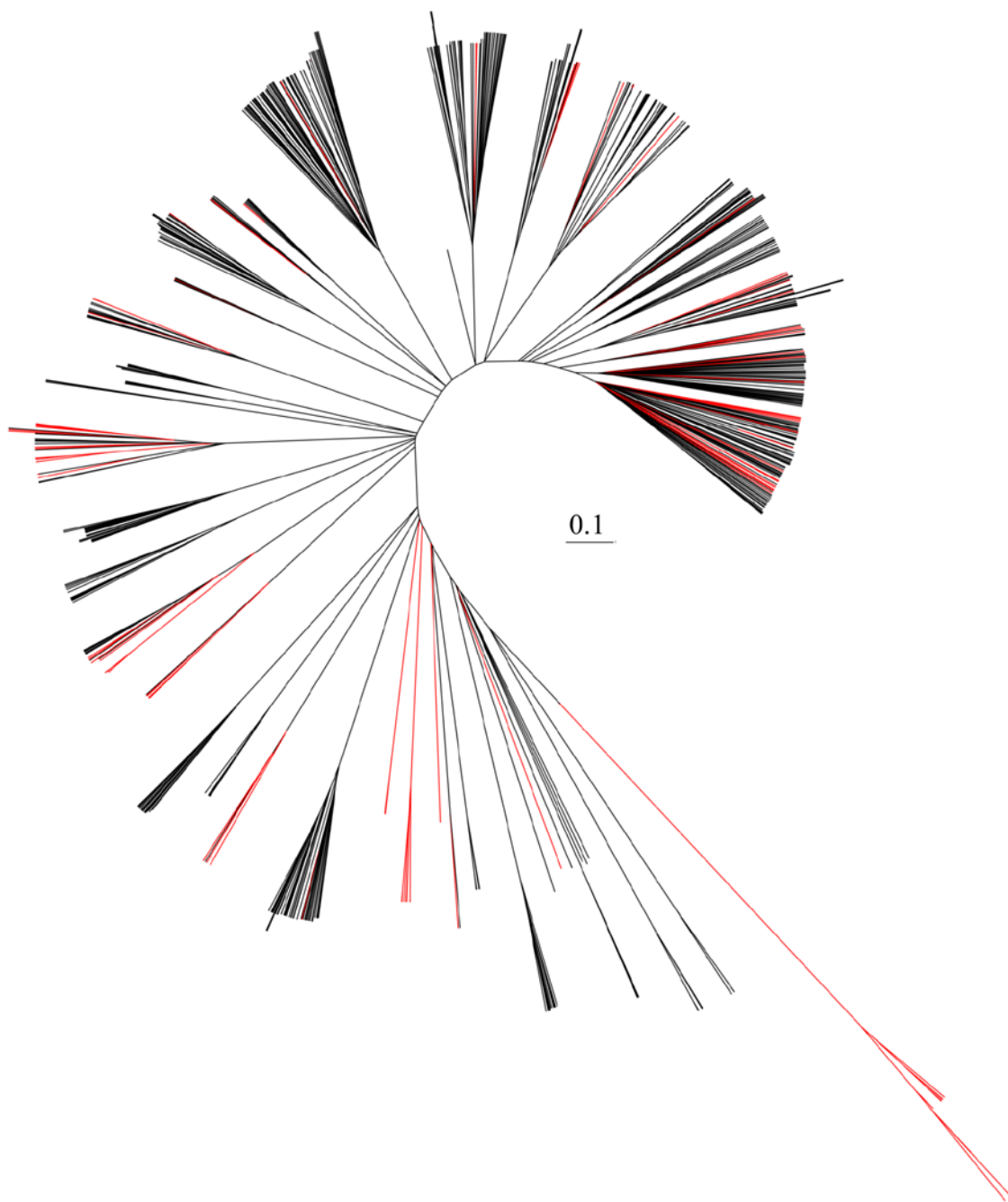


Figure 4. Terminase large subunit (LSU) protein maximum likelihood tree of Nahant Collection phages (n=264 sequences, red) with terminase LSU proteins from the ACLAMEdb of mobile elements (n = 830, black). Sequences were aligned using muscle, FastTree was used to infer an approximate-maximum likelihood phylogenetic tree, the tree was visualized using iTOL, scale bar (Edgar, 2004a, b; Letunic and Bork, 2007, 2011; Price et al., 2009, 2010).

Table 1. General procedure for tubeless agar overlay

Step	Procedure
1	Prepare 0.4% top agar in a bottle with a stir-bar, melt the top agar in a microwave prior to use, place bottle in a beaker-waterbath, maintain with stirring on a hot plate until temperature equilibrates to 50° C as indicated on a thermometer in the water bath.
2	Prepare mixture of 100uL of overnight host culture and phage treatment of interest (for example, 10uL of a dilution series of phage stock or 10uL of an environmental phage concentrate).
3	Dispense host and phage mixture onto bottom agar in a standard petri dish (100mm diameter).
4	Dispense 2.5mL of top agar onto the bottom agar.
5	Swirl to mix.
6	Place at appropriate incubation conditions with top agar side up.

Table 2. Comparison of titers in plaque forming units (PFU) as determined by tubeless versus standard agar overlay approaches.

Phage	Morphology	Tubeless (PFU/mL)	Tube (PFU/mL)	Paired t-test	
Vibriophage 12G01 ϕ	Myovirus	$1.60 \times 10^{10} \pm 1\%$	$1.52 \times 10^{10} \pm 11\%$	$t_{0.05(2),(2)} = 0.88$	$p = 0.471$
Vibriophage Jenny	Siphovirus	$1.41 \times 10^{11} \pm 3\%$	$1.57 \times 10^{11} \pm 3\%$	$t_{0.05(2),(2)} = -6.80$	$p = 0.021$
Vibriophage Al	Podovirus	$1.39 \times 10^{11} \pm 9\%$	$1.52 \times 10^{11} \pm 11\%$	$t_{0.05(2),(2)} = -1.53$	$p = 0.266$

Table 3. General procedure for serial purification of phage by streaking-for-singles into molten agar overlay

Step	Procedure
1	Insert a sterile toothpick into the plaque, or liquid phage source, to be purified.
2	Swirl the toothpick in a small area of a still-molten agar overlay of 100uL of overnight host culture.
3	Use a second toothpick to make three separate strokes in the same direction through the top agar, each time passing through the area where the first toothpick was touched.
4	Use a third toothpick to make a repeating Z-stroke through the top agar, first passing once through the streaks from the second toothpick.
5	Place at appropriate incubation conditions with top agar side up.
6	Once plaques appear on the first plate, use this plate as a source, insert the toothpick into a single plaque and repeat from step 2.
7	Single plaques arising on the third serial purification plate can be picked and stored, or used to generate liquid or plate lysates of the purified bacteriophage.

Table 4. General procedure for archiving and storage of plaques of novel phage isolates.

Step	Procedure
1	Over the course of the incubation period, mark new plaques at each time point by circling with a unique color pen for that time point. At the end of the incubation period (several days to several weeks) number all plaques and catalog them, including relevant host, sample, and morphotype data (see example in Table 5).
2	Fill a polypropylene 96-well PCR plate with 200uL aliquots of 0.2uM filtered media, collect agar plugs of plaques using a 1mL barrier pipette tip and eject into the media, skipping one well between each sample to minimize potential for cross-contamination, for a final count of 48 phage plugs per plate.
3	Soak plaque plugs at 4°C for several hours to allow elution of phage particles into the media.
4	After soaking, centrifuge 96-well plates at 2,000rcf for 3 minutes before proceeding to the next step.
5	Process plates by transferring 150uL of eluate from each well to a 0.2um filtration plate (Millipore, Multiscreen HTS GV 0.22um Filter Plate, cat# MSGVS2210) and then gently filter under vacuum to remove bacteria. Store the cell-free filtrate containing eluted phage particles from each plaque plug at 4°C.
6	Add 50uL of 50% glycerol to the residual ~50uL of the plug elution, which may still contain the original agar plug and store at -20°C.

Table 5. Example catalog sheet for tracking of large numbers of plaques.

Plaque ID	96-well Plate ID	Well ID	Host Strain ID	Phage Source	Notebook Page	Working # on plate	Definitely a Plaque? (Y/N)	Center Clear or Turbid? (C/T)	Halo? (Y/N)	Growback Colonies? (Y/N)	Irregular Margin? (Y/N)	Satellites? (Y/N)	Rays? (Y/N)	Size (mm)	Day Noted
10N.p0037.1A	10N.p0037.	1A	10N.286.52.C6	10N.286.40/41/42	K2.X.48	1	Y	C	Y	N	Y	N	N	12	1
10N.p0037.1C	10N.p0037.	1C	10N.286.52.C6	10N.286.40/41/42	K2.X.48	2	Y	C	Y	N	N	N	N	15	1
10N.p0037.1E	10N.p0037.	1E	10N.286.52.C7	10N.286.40/41/42	K2.X.48	1	N								
10N.p0037.1G	10N.p0037.	1G	10N.286.52.C12	10N.286.40/41/42	K2.X.48	1	Y	C	N	N	N	N	N	1	1
10N.p0037.2B	10N.p0037.	2B	10N.286.52.C12	10N.286.40/41/42	K2.X.48	2	Y	C	N	N	N	N	N	2	7
10N.p0037.2D	10N.p0037.	2D	10N.286.52.E3	10N.286.40/41/42	K2.X.48	1	Y	C	Y	N	Y	Y	N	7	1
10N.p0037.2F	10N.p0037.	2F	10N.286.52.E3	10N.286.40/41/42	K2.X.48	2	Y	C	Y	N	Y	Y	N	10	1
10N.p0037.2H	10N.p0037.	2H	10N.286.52.E10	10N.286.40/41/42	K2.X.48	1	Y	T	Y	Y	N	N	N	16	1
10N.p0037.3A	10N.p0037.	3A	10N.286.52.E10	10N.286.40/41/42	K2.X.48	2	Y	C	Y	Y	N	N	N	13	7
10N.p0037.3C	10N.p0037.	3C	10N.286.52.E10	10N.286.40/41/42	K2.X.48	3	Y	T	Y	Y	N	N	N	13	7
10N.p0037.3E	10N.p0037.	3E	10N.286.52.E10	10N.286.40/41/42	K2.X.48	4	Y	C	Y	Y	N	N	N	20	7
10N.p0037.3G	10N.p0037.	3G	10N.286.52.E10	10N.286.40/41/42	K2.X.48	5	Y	T	Y	N	N	N	N	11	7
10N.p0037.4B	10N.p0037.	4B	10N.286.52.E10	10N.286.40/41/42	K2.X.48	6	Y	T	N	N	N	N	N	3	13
10N.p0037.4D	10N.p0037.	4D	10N.286.52.E12	10N.286.40/41/42	K2.X.48	1	Y	C	Y	Y	N	N	N	15	1
10N.p0037.4F	10N.p0037.	4F	10N.286.52.E12	10N.286.40/41/42	K2.X.48	2	Y	C	Y	Y	N	N	N	10	1
10N.p0037.4H	10N.p0037.	4H	10N.286.52.E12	10N.286.40/41/42	K2.X.48	3	Y	C	Y	Y	N	N	N	10	1
10N.p0037.5A	10N.p0037.	5A	10N.286.52.E12	10N.286.40/41/42	K2.X.48	4	Y	C	Y	Y	N	N	N	8	1
10N.p0037.5C	10N.p0037.	5C	10N.286.52.F6	10N.286.40/41/42	K2.X.48	1	Y	C	Y	Y	N	N	N	45	1
10N.p0037.5E	10N.p0037.	5E	10N.286.52.F9	10N.286.40/41/42	K2.X.48	1	Y	T	Y	Y	N	N	N	7	1
10N.p0037.5G	10N.p0037.	5G	10N.286.52.F10	10N.286.40/41/42	K2.X.48	1	Y	T	N	N	N	N	N	10	7
10N.p0037.6B	10N.p0037.	6B	10N.286.52.F10	10N.286.40/41/42	K2.X.48	2	Y	T	N	N	N	N	N	11	7
10N.p0037.6D	10N.p0037.	6D	10N.286.52.F10	10N.286.40/41/42	K2.X.48	3	Y	T	N	N	N	N	N	12	7
10N.p0037.6F	10N.p0037.	6F	10N.286.52.F10	10N.286.40/41/42	K2.X.48	4	Y	T	N	N	N	N	N	9	7
10N.p0037.6H	10N.p0037.	6H	10N.286.55.A10	10N.286.40/41/42	K2.X.49	1	Y	C	Y	Y	N	N	N	22	1
10N.p0037.7A	10N.p0037.	7A	10N.286.55.A2	10N.286.40/41/42	K2.X.49	1	Y	C	N	Y	N	N	N	1	13
10N.p0037.7C	10N.p0037.	7C	10N.286.55.A2	10N.286.40/41/42	K2.X.49	2	Y	C	N	Y	N	N	N	1	7
10N.p0037.7E	10N.p0037.	7E	10N.286.55.A2	10N.286.40/41/42	K2.X.49	3	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.7G	10N.p0037.	7G	10N.286.55.A2	10N.286.40/41/42	K2.X.49	4	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.8B	10N.p0037.	8B	10N.286.55.A2	10N.286.40/41/42	K2.X.49	5	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.8D	10N.p0037.	8D	10N.286.55.A2	10N.286.40/41/42	K2.X.49	6	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.8F	10N.p0037.	8F	10N.286.55.A2	10N.286.40/41/42	K2.X.49	7	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.8H	10N.p0037.	8H	10N.286.55.A2	10N.286.40/41/42	K2.X.49	8	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.9A	10N.p0037.	9A	10N.286.55.A2	10N.286.40/41/42	K2.X.49	9	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.9C	10N.p0037.	9C	10N.286.55.A2	10N.286.40/41/42	K2.X.49	10	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.9E	10N.p0037.	9E	10N.286.55.A2	10N.286.40/41/42	K2.X.49	11	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.9G	10N.p0037.	9G	10N.286.55.A2	10N.286.40/41/42	K2.X.49	12	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.10B	10N.p0037.	10B	10N.286.55.A2	10N.286.40/41/42	K2.X.49	13	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.10D	10N.p0037.	10D	10N.286.55.A2	10N.286.40/41/42	K2.X.49	14	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.10F	10N.p0037.	10F	10N.286.55.A2	10N.286.40/41/42	K2.X.49	15	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.10H	10N.p0037.	10H	10N.286.55.A2	10N.286.40/41/42	K2.X.49	16	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.11A	10N.p0037.	11A	10N.286.55.A2	10N.286.40/41/42	K2.X.49	17	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.11C	10N.p0037.	11C	10N.286.55.A2	10N.286.40/41/42	K2.X.49	18	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.11E	10N.p0037.	11E	10N.286.55.A2	10N.286.40/41/42	K2.X.49	19	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.11G	10N.p0037.	11G	10N.286.55.A2	10N.286.40/41/42	K2.X.49	20	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.12B	10N.p0037.	12B	10N.286.55.A2	10N.286.40/41/42	K2.X.49	21	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.12D	10N.p0037.	12D	10N.286.55.A2	10N.286.40/41/42	K2.X.49	22	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.12F	10N.p0037.	12F	10N.286.55.A2	10N.286.40/41/42	K2.X.49	23	Y	C	N	Y	N	N	N	<1-1	1
10N.p0037.12H	10N.p0037.	12H	10N.286.55.A2	10N.286.40/41/42	K2.X.49	24	Y	C	N	Y	N	N	N	<1-1	1

Table 6. General procedure for preparation of high titer phage stocks.

Step	Procedure
1	Select a single discrete plaque in an agar overlay and use a 1mL pipette tip (not attached to a pipetter) to collect a coring through the top and bottom agar at that site. Eject into 250uL of host-appropriate media or buffer in a microcentrifuge tube and store at 4°C overnight to elute the phages from the agar plug. Also inoculate a small host culture in preparation for the next day.
2	Allow plaque eluate to come to room temperature and then spike in 20uL of overnight host culture to create a primary small scale enrichment and grow with shaking for several hours (or overnight if there is some <i>a priori</i> expectation of a long infection cycle or slow host growth).
3	Centrifuge the primary lysate at 5,000g for 5 minutes to pellet cells, transfer supernatant to fresh tube and repeat once more.
4	Prepare a 10x dilution series in host-appropriate media and pipette 5uL drop spots onto a fresh host agar overlay lawn.
5	Once plaques have developed in the drop spot series identify the dilution that gives confluent lysis and use the ratio of drop spot size to 150mm plate surface area to calculate the appropriate volume of primary lysate to use to achieve confluent lysis on the large plate. Several smaller standard size petri dishes can also be used in place of a single large petri dish and appropriate volumes per plate should be calculated accordingly.
6	Prepare an agar overlay incorporating the volume of phage determined in step 5, as well as a negative containing no phage that serves as a useful reference in assessing plaque development.
7	Harvest the top agar containing phages once the plaques have reached near-confluent lysis of the host lawn. Add 25mL of host appropriate media, or buffer, shred the top agar layer into very thin ribbons using a sterile dowel or pipette tip in hand, and collect the media and shredded top agar into a 50mL centrifuge tube. Hold at 4°C overnight to allow phages to elute from top agar.
8	Centrifuge harvested lysates at 5,000xg for 20min, syringe filter using 0.2um barrel filter (Sterivex-GP Filter Unit, Millipore), and store at 4°C. For long term storage it is recommended that 25% (final volume) glycerol stocks be prepared for storage at -20°C and -80°C, as well as of raw lysate at 4°C.

Table 7. General procedure for high-throughput spot test for host range.

Step	Procedure
1	Array phage stocks: Array 31 phage lysates, and 1 negative media control, into a 2mL 96-well culture block. For each sample, aliquot 3 replicates, distributing randomly into wells into each of three sections of the block (columns 1-4, 5-8, 9-12). Record well location of replicate.
2	Array working phage: Using a multichannel pipette, if available, transfer 150uL from each well in the deep well block to a shallow 96-well microtiter plate; prepare 4 replicate microtiter plates.
3	Pour agar overlays of host lawns: Transfer 250uL of overnight host culture and 7mL of molten top agar onto bottom agar in large petri dishes. Prepare all host lawns before starting phage application.
4	Microwave steam sterilize 96-tip blotters in batches of 6: Fill a large beaker with tap water, rinse all blotters (BelArt, Bel-blotter 96-tip replicator, cat # 378760002) vigorously, transfer to microwave steam sterilizer (Tommee Tippee, Microwave Steam Sterilizer). Microwave for 6 minutes, allow to rest for 5 minutes, transfer to PCR hood and remove sterilizer top to allow blotters to cool. For >50 host lawns use 2 sets of 6 blotters to minimize wait time and ensure sufficient pace to complete all transfers within 5-6 hours of pouring host agar overlays. Add bleach to rinse water after each batch and dispose, refilling with fresh water for the next round.
5	Apply phage to host lawn: Tap blotters on bench edge to eliminate excess moisture, set blotter into microtiter plate. Lift blotter from microtiter plate and observe to ensure that all wells have collected phage lysate by capillary action. Transfer blotter to agar overlay, gently setting down on the surface of the lawn. Work in batches of 6 host lawns, placing a blotter on each lawn and continuing to the next host lawn. All blotters should remain in place on the lawns for at least 1 minute. Lift blotters gently from lawns, transfer to paper towel to allow residual lysate to be absorbed and then sterilize as describe above. It is cautioned that blotters must be pre-treated before initial use to ensure retention of sufficient lysate by capillary action. Pre-treat by conducting multiple cycles of sterilization and blotting with phage-free media, observing for increased retention of media. Pretreatment should be conducted in advance as it may take many cyclings of blotters to achieve sufficient capillary retention.

Chapter 3

The Nahant Collection: Content, Origins, and Nomenclature

ABSTRACT

The Nahant Collection includes a variety of products derived from a consecutive 93-day time series study sampling of littoral zone marine microbes. The aim of this document is to provide a reference document for the Nahant Collection, with detailed accounting of: the collection and processing methods, the samples and isolates produced, and the nomenclature; and to offer considerations for those intending to use similar approaches in the future. This chapter also contains detailed descriptions of all relevant methods for Chapter 4. The information here is presented in three sections: I. Daily Field Collections, II. Large Scale *Vibrio* Isolations, and III. Phage Isolation (Large Scale Plating) & Plaque Archival.

INTRODUCTION

The Nahant Time Series was conducted over a period of 93 consecutive days in order to provide a comprehensive, highly temporally resolved, sampling of marine microbial communities. This littoral zone time-series study captured information about abiotic conditions, microbial populations, and higher order taxa in the littoral ocean at Nahant, MA, on a daily basis over a period of three months. The sampling period between July 23 and October 23, 2010, spanned a seasonal transition and was superimposed by several storm events, North Atlantic hurricanes, and possible intrusion of cold water masses. Daily measures of air and water temperature, and salinity, were recorded on site and were supplemented with external wind, rainfall, and tide datasets. Daily samples were collected for analysis of nutrients, including ammonium, nitrate, silicate, and phosphate. Daily archives of the biological diversity include DNA and live cells of the microbial community from unfractionated, 63 μ m, 5 μ m, 1 μ m, and 0.2 μ m size fractions, and live virus concentrates. Large scale isolations of strains at each of 3 main time points spanning the time series provide a collection of nearly 3,500 purified isolates, mainly *Vibrio*. Direct plating of viral concentrates from the same sampling days onto >1,300 bacterial isolates yielded quantitative representations of co-occurring phage predators. A collection of ~2,000 morphologically-described plaques picked from these plates is available in archive. A subset of ~250 of these plaques was purified and genome sequenced and characterized for host range with respect to ~300 hosts. It is hoped that data collected in this study will enable analyses that integrate the role of physical, chemical, and biotic forcing in assembling and structuring marine microbial communities.

Section I. Daily Field Collections

Duration. Samples were collected in the littoral zone at Canoe Cove, Nahant, MA (42.41975, -70.907114) [Figure 1], every day between July 23, 2010 (ordinal day 204) and October 23, 2010 (ordinal day 296).

Sampling Day. Collection at Nahant generally occurred between 1000 and 1100 in the morning. Sample processing in the lab generally began by 1230 and finished around 1900. At the close of each day all materials were cleaned, autoclaved, and set-up for the following day.

Sampling. Samples were collected by wading into the littoral zone to a depth of several feet. The absolute location of sampling day varied each day dependent on tidal height, given the low slope of the beach this was quite variable.

Stations Change in Sampling Regime: Prior to Day 259 triplicate samples for each sample type were collected together at an undefined location, though generally closest to the location designated as station 2 starting on Day 259. Collections for different sample types collected in this way were not necessarily collected in the same place if collected by different workers. From Day 259 forward replicates for each sample type were collected at 3 defined stations (Fig. 2). *Station 1:* was close to a rocky outcropping that was covered in macroalgae, especially *Fucus* sp. and *Ascophyllum* sp.. If there is a signal associated with live attached macroalgae it would be expected to be strongest in samples collected at this station. *Station 2:* was the middle station and was determined as a straight walk from the flag pole on the Nahant campus to the water. *Station 3:* was near the middle of Canoe Cove and was determined as a straight walk from the fence dividing Nahant property from the public beach down to the water (Figure 2).

Temperature. Air and water temperature were recorded using analog & digital thermometers. Luke Miller (Postdoc, NEU) shared an independent dataset containing detailed submerged temperature records he collected on the Boston Harbor side of Nahant. Though these have not been used in analyses to date they are available.

Salinity. Salinity was measured using a refractometer post-study on stored samples due to equipment failures during the time course. For future studies it is recommended that a hand-held refractometer be used in the lab on samples collected in the field to prevent equipment damage on site.

Additional Datasets. Additional datasets useful in the context of this study are available from the National Oceanographic and Atmospheric Administration (NOAA). Buoy station 444013 is 16 nautical miles east of Boston and available datasets for this station include: ocean current data, continuous winds data, standard meteorological data, ocean data (including chlorophyll concentration), spectral wave density data, and spectral wave direction data. Data from these datasets corresponding to the dates of the study were originally downloaded from (www.ndbc.noaa.gov/station_page.php?station=44013) and will be available in (for data location see Table 1).

General Note on Sample Nomenclature. All field samples associated with the Nahant Collection have a shared nomenclature: 10N.xxx.yy, based on year (10, 2010), location ID (N, Nahant), the ordinal day of collection ID (xxx), and a sample type ID (yy) as defined in the daily data sheet (example of daily data sheet, Figure 3).

General Note on Size Fractionation & Vibrio Isolation: The methods employed for size fractionation of water samples throughout this study are intended to be identical with those previously used for large scale isolations of *Vibrio* in the Polz Lab and as taught to this author by Sarah Pacocha Preheim and originally developed for (Hunt et al., 2008). Minor modifications are included to increase throughput, for example, use of peristaltic pump filtration and Sterivex (Sterivex, SVGP01050, Millipore) barrel filters for the 0.2um fraction rather than cup and frit tower systems. Initial attempts to employ in-line arrays of syringe filter holders for size-fractionation, rather than cup and frit tower systems, proved logistically undesirable at high throughput due to constant leaks at all junctions following required autoclaving.

Unfractionated DNA & Phage Samples. Unfractionated DNA and phage concentrates derive from the same water samples and were processed sequentially – first for collection of unfractionated DNA (sample numbers 37, 38, and 39), then for concentration of phage (sample numbers 40, 41, and 42). Each day three ~4L samples were collected in screw-cap polypropylene bottles. The bottles were marked with a line at the 4L mark at the beginning of the study. The night before sampling the bottles were rinsed 3x with tap water and 3x with RO water. For collection the bottles were always rinsed 3x with seawater just prior to collection, then the bottles were submerged and filled to above the 4L line, and excess sample was poured off. Upon return to the lab the water from each of the 4L samples was filtered by peristaltic pump

(setting ~5) from the collection container through a 0.2um Sterivex filter into a 4L collection bottle. All collection bottles were always cleaned as follows: washed with soapy water, rinsed 3x with tap water, 3x with RO water, 3x with milliQ water, and then autoclaved. Some samples required multiple Sterivex filters to be completely filtered, as there were no direct measurements of salinity collected during the study the number of Sterivex filters required to process the unfractionated samples may be considered the best available proxy. A syringe was used to attempt to blow-out residual liquid before storage of the Sterivex filters at -20°C. The Sterivex filters are the “Unfractionated DNA” samples collected during the Nahant study. Antonio Martin Platero has extracted DNA from all of these Unfractionated DNA Sterivex filters (sample IDs 37, 38, 39) and prepared, sequenced, and analyzed 16S and 18S amplicon libraries for all of them. The Sterivex filters have been thrown away and the extracted DNA has been stored.

The water collected following filtration through the 0.2um Sterivex was spiked with 200uL of a control phage lysate in M17+++ broth (DT1 phage of *Streptococcus thermophilus* SMQ-301, both from the Felix D’Herelle Reference Center for Bacterial Viruses). Following addition of control phages the bottles were capped and gently inverted 10x to mix. The water was then spiked with 400uL of FeCl₃ solution, as per [Appendix 1, (John et al., 2011)], and allowed to incubate at room temperature for at least 1 hour. The 4L water samples were then filtered onto 90mm 0.2um polycarbonate filters (Isopore, GTTP09030, Millipore). Samples were filtered by vacuum pump at a pressure of setting of ~10” Hg using a glass cup/frit/flask set-up with an approximately 1L capacity. Processing each sample required several re-fills of the cup and emptyings of the receptacle flask. The bottle was gently inverted or swirled prior to each fill of the cup and when the flask was nearly full the cup was lifted and the water from the receptacle flask poured into a graduated cylinder for measurement. Following complete filtration of all of the sample the filter was folded into quarters under vacuum and inserted into a 7mL borosilicate glass vial. The borosilicate glass vials were rinsed with milliQ and autoclaved prior to use. A volume of 4mL of oxalate solution was then added to the vial and the sample allowed to dissolve at room temperature for at least 30 minutes before transfer to storage at 4°C. The oxalate solution was prepared fresh approximately every 7 days from solutions of MgEDTA, Oxalate, and NaOH. The final oxalate solution and the component reagents were all maintained at room temperature in the dark [for recipe see Appendix 1, (John et al., 2011)]. Phage concentrate replicates for each day were pooled into a single sample in 2013. A time series direct plating using all available

pooled replicates and a set of 10 Nahant isolates was conducted in 2013 by Fatima Hussain and the resulting plaques archived.

63um Samples (for DNA &Community Glycerols) 63um samples were collected by pouring known volumes of water through a plankton net (63um Nitex Turtox Tow Net, Wild Co., 426-A38) with an Oakridge tube screwed in at the cod end. Approximately 100L were poured through the net (measured using a 10L bucket). To concentrate the retained particles the top of the net was twisted closed and the net dipped into the seawater repetitively to wash particles down the net and into the collection tube. When the tube was full the net was twisted shut just over the opening and the tube turned upside down so that the net could be pumped up and down into the tube to milk out water until the final level was even with the 30mL line marked on the tube at the beginning of the study. In this way material from 100L was concentrated down to 30mL. In the lab the samples were poured into tissue grinders (47732-450, VWR) for manual homogenization. On some days all samples were ground using the same grinder, on these days the grinders were rinsed first with autoclaved milliQ water and then with 0.2um filtered artificial seawater in between samples. For DNA, 10mL of homogenized concentrate were then filtered through a 0.2um filter (during the first days of the study a cup/frit system and a 0.2um polycarbonate filters were used and then later a change was made to using Sterivex 0.2um filter with a syringe). For community glycerols, 10mL of homogenized concentrate was filtered onto a 0.2um polycarbonate filter using a cup/frit system. The filter was then transferred to 4mL of 50:50 2x 2216:50% glycerol, final concentration of 1x 2216 and 25% glycerol, and shaken for approximately 20 minutes before freezing at -80°C.

63um Samples- Nomenclature for DNA &Community Glycerols. DNA and community samples originating from the same replicate bottle were as follows, where xxx indicates the ordinal day of collection:

Replicate 1: DNA (10N.xxx.13), Glycerol (10N.xxx.16)

Replicate 2: DNA (10N.xxx.14), Glycerol (10N.xxx.17)

Replicate 3: DNA (10N.xxx.15), Glycerol (10N.xxx.18)

DOC Samples. Samples for dissolved organic carbon were collected using 10mL syringes and a 0.2um Acrodisc filter. All 3 samples were collected using the same syringe and Acrodisc filter. The syringe was rinsed out 3x with seawater, then seawater was sucked up and dispensed into a collection vial through the 0.2um Acrodisc filter for an initial rinse of the collection tube. During rinsing care was taken not to allow contact between the sample water and the aluminum cap. A sample of 10mL of 0.2um seawater was then dispensed into the collection tube and the tube capped with aluminum foil and placed in a cooler with ice packs. The collection vials were 14mL screw cap glass culture tubes that had been acid washed. Upon return to the lab 5 drops of phosphoric acid were added to each sample and the pH checked with a pH strip to ensure pH was around 1.5. Samples were then stored at 4°C with fresh aluminum foil coverings until the end of the study, when the caps were replaced with Teflon lined caps. Each day packets of 3 collection tubes were placed in a baggie with an Acrodisc filter and a syringe for collection the following day. Blank DOC samples were collected by taking tubes into the field and upon return to the lab rinsing and filling them with milliQ resin purified water. Syringes were rinsed as for standard samples, then filled with Q water and a small amount dispensed into the collection tube 3 times for a rinse before filling with 10mL of Q water and addition of phosphoric acid as for the other samples. DOC samples were analyzed at UNH and results showed background in excess of sample values rendering all samples unusable. For future studies it is strongly recommended that preliminary collections be made and submitted for analysis to confirm that methods employed (even if standard) yield sufficient removal of background.

5um, 1um, 0.2um Samples (for DNA and Community Glycerols). Fractionated samples for DNA and community glycerols were collected by passing a single sample of water through serially smaller pore-sized filters using cup and frit filtration towers. The water sample was collected on site into an autoclaved 1L bottle pre-rinsed with 63um-prefiltered seawater on site, collection of the sample was by pouring water from a bucket through the 63um net and filling up the 1L bottle after initial pre-rinsing with 63um filtered water. Set-up of the filtration array in the lab included 1 flask and tower set-up per collection-bottle replicate, all flasks and towers are washed, rinsed in reverse osmosis purified water and autoclaved prior to use. Prior to initiating filtration, all filters for each sample (5um, 1.0um; and 0.2um for glycerols only) are put in place on their respective filtration towers. Water is poured from the collection bottle into the tower of the 5um filtration set-up and allowed to gravity filter through, filtrate is transferred by pouring

off into the 1um tower for further gravity filtration, filtrate from the 1um tower is either pumped directly through a 0.2um sterivex filter for DNA samples or transferred by pouring into the 0.2um tower where it is vacuum filtered under <10”Hg pressure. Filters for DNA are folded on the frit and transferred to microcentrifuge tubes for storage at -20°C or -80°C. To collect filters for community glycerols the same procedure was repeated, using the same series of tower set-ups for water from the same replicate bottle, and the filter was then transferred to 4mL of 50:50 2x 2216:50% glycerol, final concentration of 1x 2216 and 25% glycerol, and shaken for approximately 20 minutes before freezing at -80°C. Antonio Martin Platero has extracted DNA from all of these fractionated DNA filters (sample IDs listed below) and prepared, sequenced, and analyzed 16S and 18S amplicon libraries for all of them. The filters have been thrown away and the extracted DNA has been stored.

5um, 1um, 0.2um Samples – Nomenclature for DNA and Community Glycerols

Replicate A:

DNA-5um (10N.xxx.19), Glycerol-5um (10N.xxx.29)

DNA-1um (10N.xxx.22), Glycerol-1um (10N.xxx.31)

DNA-0.2um (10N.xxx.25), Glycerol-0.2um (10N.xxx.34)

Replicate B:

DNA-5um (10N.xxx.20), Glycerol-5um (10N.xxx.29)

DNA-1um (10N.xxx.23), Glycerol-1um (10N.xxx.32)

DNA-0.2um (10N.xxx.26), Glycerol-0.2um (10N.xxx.35)

Replicate C:

DNA-5um (10N.xxx.21), Glycerol-5um (10N.xxx.30)

DNA-1um (10N.xxx.24), Glycerol-1um (10N.xxx.33)

DNA-0.2um (10N.xxx.27), Glycerol-0.2um (10N.xxx.36)

Nutrient Samples. For nutrient sample collection and processing all bottles, filtration towers & frits, and collection flasks were acid washed before use. All further references to acid washing

imply soaking overnight in 1M HCl bath, followed by rinsing with copious amounts of milliQ water. Samples were collected by submerging a 250mL polypropylene bottle in seawater at each of the different stations, prior to collection each bottle was rinsed three times with seawater on site. Samples were stored in a cooler for transport back to lab. On haphazardly selected days blanks were processed before and after field sample processing to control for contribution of equipment to nutrient signal, and to assess carry over between samples. Lab filtration procedure, including processing of blanks was as follows. The filtration tower was assembled and rinsed again by running 100mL milliQ-HOH through the assembled apparatus. For collection of the initial equipment blank an ashed GF/F filter was placed on the frit and then 150mL of milliQ was filtered into the collection flask and partitioned by pouring into 4 acid-washed scintillation vials. For each seawater sample replicate a fresh ashed GF/F filter was placed on the frit and 150mL of seawater was poured into the filtration cup, filtrate was used first to rinse out the 4x 20mL scintillation sample storage vials and then ~15mL of sample was poured into each vial for storage at -20°C or -80°C. After each replicate was partitioned the residual filtrate was poured away and the tower cup was refilled with milliQ with the GF/F from the previous sample in place to rinse the cup and flask prior to continuing with the next sample. A new GF/F was set in place between each replicate. A final carry-over blank was collected by filtering 150mL of milliQ through a fresh ashed GF/F filter and collecting the filtrate as for samples. Ashed GF/F filters were prepared by wrapping ~10 filters together in aluminum foil pouches and baking them overnight at 450°C. Nutrients were analyzed by the Nutrient Facility at Woods Hole Oceanographic Institution. Expansion of liquid in storage bottles resulted in overflow from some stored nutrient samples, values for phosphate, nitrate, and ammonium were corrected based on correction factors as determined and provided by Hanan Karam (Karam, 2012). Correction factors for silicate were assumed to be the same as that used for the above nutrients but this has not been tested to date. For future applications the collection of a single 250mL is recommended.

Chlorophyll Samples. Samples processed for chlorophyll were collected in 1L bottles that had been washed and autoclaved. Bottles were rinsed three times in seawater just prior to sample collection. In the lab the samples were processed on the same set-up as had been used for processing nutrient samples after the nutrient samples had been processed and the tower and frit had been rinsed with milliQ water. A blank of 500mL of milliQ was collected on GF/C filter prior to processing the seawater sample. The chlorophyll sample was processed by filtering

500mL of seawater through a GF/C filter, wrapping the filter in aluminum foil and storing at -20°C during the course of the study and -80°C afterwards. For future studies -80°C storage is recommended, and it is also noted that GF/F, GF/C, and 0.2um filters are all described for use for this application in the literature and that GF/C has a larger pore size and so this would be expected to yield lower recoveries than other filters and thus perhaps not the best choice. At the time of this writing chlorophyll samples remain unprocessed and stored at -80°C.

Counts. Samples for counts were collected in 50mL falcon tubes rinsed three times with seawater before use; 48mL were collected for each sample and transferred to a 125mL bottle where they were spiked with 2mL of 0.02 filtered formalin (to eliminate virus-sized particles because of use in viral counts). Collection bottles were new and prepared by rinsing with hot water and then rinsing and soaking in milliQ water. Samples were stored in a cooler and upon return to lab were either stored at -80°C immediately (during the beginning of the study) or submerged in an ethanol-dry ice bath for rapid freezing before transfer to storage at -80°C (later in the study). At the time of this writing samples for counts remain unprocessed and stored at -80°C.

Microscopy, Field Notes, and Photos. Extensive qualitative metadata is available for the majority of the time series. Included are: A) extensive qualitative notes on weather and sampling site conditions, these are recorded on the data sheets as well as in abbreviated form in excel spreadsheets; B) extensive qualitative notes on observations of microscopy of small subsamples of raw unfractionated samples; C) daily photographs of the field site, and some movies, these begin some time into the study; and, D) daily photographs of microscopy, and some movies, these also begin some time into the study. For examples of (A) and (B) see Figure 3; for examples of (C) see Figure 4 (for data location see Table 2).

Additional Notes. It is recommended that all types of sample collection methods be pre-tested, including a complete analysis, before deployment in the field during a time course. It is recommended that options for *in situ* data loggers be explored, these would offer higher consistency and finer resolution of environmental data. For those considering future work at the site at Nahant it is noted that it is necessary to complete insurance arrangements between MIT and the Northeastern University Marine Science Center if intending to collect in areas contained within the bounds of their field station (as Sampling Stations 1 and 2 of this study are).

People Involved: *Field sampling:* Kathryn Kauffman (principal; every day), Tara Soni, Gitta Szabo, Alison Takemura, Aidong Ruan, Robert Ratzlaff, Otto Cordero, Michael Cutler, Hans Wildschutte, Nisha Vahora, Martin Polz, Jane Cote, Hong Xue, Hanan Karam, Antonio Martin Platero *Lab processing:* Kathryn Kauffman (principal; every day), Alison Takemura (principal for processing of nutrients for large portion of time series), Tara Soni, Hong Xue, Michael Cutler, Nisha Vahora, Gitta Szabo, Su Xue, Otto Cordero, Hans Wildschutte, Robert Ratzlaff; *Lab support:* Michael Cutler, John Macfarlane; *References:* Hanan Karam (Nutrients), Nisha Vahora (DOC); *Logistical Support:* Michael Cutler, Sara Phenix; *Data:* Luke Miller (Temperature, not used to date, but shared; for data location see Table 1).

Section II. Large scale *Vibrio* isolations

Bacteria Isolation. At four time points in the sampling time series (ordinal days 222, 250, 261, and 286) water samples were collected for fractionation and isolation of *Vibrio* using previously established methods (Hunt et al., 2008). Water samples were collected in the same way as for the daily size fractionation, material recovered on filters was resuspended from filters by shaking for 20 minutes, and then different volumes of resuspended cells were filtered onto 0.2um PES filters in a carrier solution of artificial seawater (40g Sigma Sea Salts, S9883; 0.2um filtered) which were then placed directly onto agar plates of MTCBS *Vibrio*- selective media [thiosulfate-citrate-bile-sucrose (TCBS Agar, Difco) prepared according to manufacturer instruction and amended with 10g NaCl per liter to 2% final w/v). Colonies were allowed to form on the filters and then 96 colonies from each replicate of each size fraction were selected from the dilution plates with the fewest numbers of colonies, for a total of 1,152 isolates per isolation day. Colonies were purified by serial passage, first onto TSB-II (Tryptic Soy Broth, 1.5% Difco Bacto Agar, amended with 15g NaCl to 2% w/v), second onto MTCBS, finally onto TSB-II again. Colonies were inoculated into 1mL of 2216 Marine Broth (Difco) in 96-well 2mL culture blocks and allowed to grow, shaking at room temperature, for 48 hours. Glycerol stocks were prepared by combining 100uL of culture with 100uL of 50% glycerol in 96-well microtiter plates and sealing with adhesive aluminum foil. Isolates are stored in replicate plates with the nomenclature indicated as described in Table 2. It is noted that the isolate collection of ordinal day 250 was

prepared without the addition of salt to the TCBS and that it is therefore a much smaller collection and likely of distinct composition, possibly enriched in *V. cholerae*.

***Hsp60* gene sequencing.** All *Vibrio* isolates used for the Large Scale Plating (LSP) test of phage susceptibility were sequenced at a single locus, the gene for the heat shock protein *Hsp60*, in order to determine their phylogenetic relationships and assign them to previously recognized populations of *Vibrio* with well-characterized ecologies. Lyse-N-Go treatments (Thermoscientific, 78882, cycled as per manufacturer instructions) of subsamples of the same overnight cultures used in the LSP were used directly as template in PCR amplification reactions. PCR reactions were prepared in 30uL volumes, with the recipe: 1uL LNG template, 3uL 10x buffer, 3uL 2mM dNTPs, 3uL 2uM hsp60-F primer, 3uL 2uM hsp60-R primer, 0.3uL NEB Taq, 16.7uL PCR-grade HOH; with hsp60-F (H279) primer sequence: 5'-GAA TTC GAI III GCI GGI GAY GGI ACI ACI AC-3', and hsp60-R (H280) primer sequence: 5'-CGC GGG ATC CYK IYK ITC ICC RAA ICC IGG IGC YTT-3' [ADD goh1996hsp60]. PCR thermocycling conditions were as follows: initial denaturation at 94°C for 2 min; 35 cycles of 94°C for 1 min, 37C for 1 min, 72C for 1 min; final annealing at 72C for 6 min; hold at 10C. PCR products were cleaned up by isopropyl alcohol (IPA) precipitation as follows: addition of 100uL 75% IPA to 30uL PCR reaction product, gentle inversion mixing followed by 25 min incubation at RT, 30 min centrifugation at 2800rcf, addition of 50uL 70% IPA with gentle inversion wash, centrifugation at 2000rcf, inversion on paper towels to remove IPA, 10 min centrifugation at 700rcf, air drying in PCR hood for 30 min, resuspension in 30uL PCR HOH. PCR products were Sanger sequenced (Genewiz, Inc.) using hsp60R primer, with the recipe: 5uL of 5uM hsp60R primer, 7uL nuclease free water, 3uL DNA template. *Hsp60* sequences were aligned to the 427 base hsp60 sequence previously published for *Vibrio* 1S_84 and trimmed to 427 bases (Geneious). It is noted that *Hsp60* amplicons from *Vibrio* isolates often include additional off-target smaller bands, the presence of such sequence may have contributed to the difficulty experienced in achieving good sequence for many isolates. It is highly recommended that, where possible, both primers be used for sequencing. Also, a small number of *Hsp60* sequences were recovered using whole-genome sequence of isolates that became available during this study and it was often found that there were discrepancies at 2 particular sites between the end of already-trimmed Sanger sequence and genome sequence that was detected for multiple isolates. The cause of this is not clear though to allow inclusion of these additional

sequences the *Hsp60* sequences were trimmed to the 422bp region for which there was never disagreement between any Sanger and genome sequenced versions. *Hsp60* sequences trimmed to 422 were mapped to previously described populations by building a single tree incorporating 1,271 Nahant Collection strains, 1,025 strains from the Hunt Collection (Hunt et al., 2008) and 935 strains from the Szabo Collection (Szabo et al., 2012), sequences were aligned using Muscle (Edgar, 2004a, b), aligned using FastTree (Price et al., 2009, 2010), and visualized using iTol (Letunic and Bork, 2007, 2011). Where there were discrepancies in naming, those used were as in (Preheim et al., 2011). This mapping showed that the strains recovered in the Nahant Collection were representative of the previously recovered populations (Figure 5).

Additional Notes. Subsequent work with strains in this collection have shown some of them to appear mixed, in some cases this is likely slight contamination from multiple uses of glycerol stocks, in other cases this may be due to true ties between strains that cause them to often be co-isolated, and in some cases this may be due to phase variation that gives rise to two colony types (flat clear and thick creamy) and which has often been observed in phage resistant versions of strains and may thus be common. In general workers are cautioned to confirm the identity of strains by sequencing of *Hsp60*, or another marker gene, followed by confirmation against the expected Sanger sequence or full genome. In cases where isolates are no longer recoverable or lost it is noted here that isolates from the Nahant Collection have been provided to other labs and may thus still be available if needed; to date: Hans Wildschutte (full collections for days 222, 261, and 286), Otto Cordero (partial set), and Frederique LeRoux (partial set).

People Involved: *Isolations & serial purifications of strains:* Many members of the Polz lab.
Hsp60 gene sequencing: Michael Cutler

Section III. Phage Isolation (Large Scale Plating) & Plaque Archival

More than 400 *Vibrio* isolates from each of the three days of the Nahant Collection were exposed individually to 15mL seawater equivalents of concentrate of co-occurring phage in agar overlays in an experiment hereafter referred to as the Large Scale Plating (LSP). These assays provided an estimate of the concentration of plaque-forming phages infecting each *Vibrio* strain. Plaques formed during this assay were isolated and stored independently, providing a source of viable phage for subsequent amplification and study. For a more general description and discussion of

the methods employed for working with phages see Chapter 2.

Strain Selection. Large volume (20L) concentrates collected on the three isolation days did not contain viable phages, as determined by assaying for presence of the spiked control phage DT1 on its *Streptococcus thermophilus* host, SMQ-301. The failure to recover viable phage may be attributable to elevated, and unknown, pressures generated during filtration of the spiked 20L sample through a 142mm polycarbonate flat filter. Future attempts to concentrate large volumes should employ equipment allowing for regulation of pressure as outlined in [Appendix 1, (John et al., 2011)]. The phages used in the LSP were therefore derived from pooling of the three replicate 4L-concentrates collected for each of the days for which strains were isolated (222, 261, and 286). Approximately 480 *Vibrio* strains from each ordinal day isolate collection were assayed for susceptibility to co-occurring phage in 15uL of phage concentrate (equivalent to 15mL of seawater). The ~1440 *Vibrio* strains assayed were divided as follows: 480 strains per ordinal day collection; for each day, 120 strains were tested for each of the 4 fractionation size classes (0.2um, 1.0um, 5.0um, 63um); within each size fraction 60 strains were taken from each of 2 replicate plates for that size fraction; and within each replicate plate all strains from rows A,B,C,E, and F were selected (K2.X.9).

Agar Overlay Plating. Strains were prepared for agar overlay plating by streaking out from glycerol stocks onto 2216 Marine Broth agar plates with 1.5% Bacto Agar (Difco), and allowed to grow for 2 days at room temperature. Strains were then inoculated into 1mL of 2216 Marine Broth in a 2 mL volume 96-well culture block and incubated 24 hours at room temperature shaking at 275rpm on a DS500E orbital shaker (VWR). Immediately prior to use in direct plating the OD600 was measured in 96-well microtiter plates and then subsamples were taken for Lyse-N-Go processing for DNA (10uL culture, 10uL LNG) and for preparation of new glycerols for storage (“D” glycerols, 100uL overnight culture, 100uL 50% glycerol). Phage concentrates were prepared for plating by pooling 1.2mL from each of the concentrate replicates into a 7mL borosilicate scintillation vial. Cultures were transferred from overnight culture blocks to 96-well PCR plates in 100uL volume and 15uL of pooled phage concentrate was added to cultures one row at a time, with each row plated in agar overlay before adding phage concentrate to the next row of *Vibrio* cultures. Mixed samples of 100uL *Vibrio* overnight and 15uL pooled phage concentrate were transferred to the surface of bottom agar plates (2216 Marine Broth, 1% Bacto

Agar, 5% glycerol, 125mL/L of chitin supplement [40g/L coarsely ground chitin, autoclaved, 0.2um filtered]). A 2.5mL volume of ~52C molten top agar (2216 Marine Broth, 0.4% Bacto Agar, 5% glycerol) was added to the surface of the bottom agar and swirled around to incorporate and evenly disperse the mixed *Vibrio* and phage sample into an agar overlay lawn. Agar overlay lawns were held at room temperature for 14-16 days and observed for plaque formation. Glycerol was incorporated into this assay to facilitate detection of plaques (Santos et al., 2009). Chitin supplement was incorporated into this assay to facilitate detection of phages interacting with receptors upregulated in response to chitin degradation products. A variety of preliminary tests exploring potential optimizations to agar compositions for direct plating indicated that the addition of chitin did not negatively impact recovery of plaques with control phage strains tested.

Negative Controls: It is noted here that two sets of important negative controls were not included due to failure to appreciate the potential significance of autoinduction. Autoinduction is the spontaneous formation of plaques on host lawns to which no external phage has been added; the potential for prophages to form plaques on host lawns is, in and of itself, an intriguing phenomenon but is unwelcome when attempting to assay for phages in environmental samples. For future studies it is strongly recommended that the following negative controls be included: 1) plating of host lawn alone, and 2) plating with oxalate resuspension of iron chloride precipitated phage-free sample, perhaps using 0.2um filtered and then autoclaved seawater from the same sample. In the case of (1) this allows a control for culture- or plating-condition related stimulation of autoinduction, in the case of (2) this allows a control for test of iron-oxalate chemistry related stimulation of autoinduction. Strains known to be autoinducers are noted in the catalog of developed phage-host systems (Table 3). Concern regarding the impacts of prophage induction on the findings of this study was addressed by conducting searches for sequence similarity between sequenced phage genomes and available host genomes. Sequenced phage genomes would also be expected to be present in the host genomes if they represent induced prophages and no evidence for this was found.

Plaque Enumeration and Storage After approximately 2 weeks, plaques on agar overlay lawns were cataloged and described with respect to plaque morphology and plaques were picked for storage. All plaques were archived from plates containing less than ~25 plaques, on plates with

larger numbers of plaques a random subsample of plaques from each distinct morphology were archived. Plaques were archived as follows. All plaques on a plate were numbered and then described for entry into a plaque morphology spreadsheet (example as in Chapter 2, for data location see Table 1). A polypropylene 96-well PCR plate was filled with 200uL aliquots of 0.2um filtered 2216 Marine Broth, agar plugs were collected from plates using a 1mL barrier pipette tip and ejected into the 2216 Marine Broth, skipping one well between each sample to minimize potential for cross-contamination, for a final count of 48 phage plugs per plate. Plaque plugs were soaked at 4°C for several hours to allow elution of phage particles into the media. After soaking, 96-well plates were centrifuged at 2,000rcf for 3 minutes before proceeding to the next step. Plug soaks were then processed for two independent storage treatments. For storage at 4°C, plates were processed by transferring 150uL of eluate from each well to a 0.2um filtration plate (Millipore, Multiscreen HTS GV 0.22um Filter Plate, cat# MSGVS2210) and then gently filtered under vacuum to remove bacteria, the cell-free filtrate containing eluted phage particles from each plaque plug were stored at 4°C. For storage at -20°C, 50uL of 50% glycerol was added to the residual ~50uL of the plug elution, often still containing the agar plug. In this way all plaques were characterized and many plaques from each strain were archived in two independent sets of conditions (K2.X.32-34). The plates containing archived plaques are labeled with the following nomenclature: 10N.p0001, where 10N indicates association with the 2010 Nahant Collection, “p” indicates phage sample, and 0001 indicates plate number. Phages can be identified as wells within plates and thus have IDs of the following form: 10N.p0001.A1. All phage host systems selected for further development derive from either the 4°C or -20°C plaque archive and thus have are associated with a particular 10N.pxxxx.xx ID of origin in addition to later working IDs.

Section IV. Purification and Sequencing of Selected Phages

A subset of plaques detected and archived during the Large Scale Plating were selected for further development. Usually this was a single phage for each of the *Vibrio* hosts on which plaques were observed during the LSP, this attempts to purify 295 phage-host pairs. Phages were serially purified and high-titer lysates were generated to serve as stocks for future culture work and for genome sequencing with Illumina (see below).

Serial Purification & Lysate Preparation. Phages were selected for serial purification by

random number generated selection of one archived LSP plaque isolate for each LSP plaque-positive host strain. Minor details of the purification and lysate preparation varied across samples but were largely as follows. Phages were serially purified, generally for 3 passages beyond initial plaque recovery from archives, using re-streaking of isolated plaques. The initial phage inocula were derived primarily from -20°C LSP plaque archives, and secondarily from 4°C archives when primary attempts with -20°C stocks failed to produce plaques. Agar overlay lawns were prepared by aliquoting 100uL of host overnight culture (4mL 2216 Marine Broth, colony inoculum from streak on 2216 Marine Broth with 1.5% Bacto Agar, shake overnight at RT at 250rpm on DS500E orbital shaker [VWR]) onto a standard size bottom agar plate (2216 Marine Broth, 1% Bacto Agar, 5% glycerol, 125mL/L of chitin supplement [40g/L coarsely ground chitin, autoclaved, 0.2um filtered]), adding 2.5mL of molten ~52C top agar (2216 Marine Broth, 0.4% Bacto Agar, 5% glycerol), swirling to disperse the host into the top agar and form a lawn, and streaking-in phage inoculum with a toothpick either from the plaque archive or from isolated plaques of the previous step in serial purification. Following plaque formation on the third serial passage plate plaque plugs were picked using barrier tip 1mL pipettes and ejected into 250uL of 2216 Marine Broth to elute overnight at 4°C. Plaque eluates were spiked with 20uL of host culture and grown with shaking for several hours to generate a primary small-scale lysate. Small-scale primary lysates were centrifuged to pellet cells and then titered by drop spot assay to estimate optimal inoculum volume to achieve confluent lysis in a 150mm agar overlay plate lysate. Plate lysates were generated by mixing 250uL of overnight host culture with primary lysate and plating in 7.5mL using standard agar overlay conditions as previously described. After development of confluent lysis of lawns as compared against negative control without phage addition, the lysates were harvested as follows: addition of 25mL of 2216 Marine Broth, shredding of the agar overlay with a dowel, collection of the broth and top agar. Freshly harvested lysates were stored at 4°C overnight to allow for elution of phage particles, the following day lysates were centrifuged at 5,000xg for 20 min and the supernatant filtered through a 0.2um sterivx filter into a 50mL falcon tube and stored at 4°C.

DNA Extraction. For DNA extraction approximately 18mL of phage lysate was concentrated using a 30kD centrifugal filtration device (Millipore, Amicon Ultra Centrifugal Filters, Ultracel 30K, UFC903024) and washed with 1:100 2216 to reduce salt concentrations inhibitory to downstream nuclease treatments. Concentrates were brought to approximately 500uL using

1:100 diluted 2216 Marine Broth and then treated with DNase I and RNase A for 65 minutes at 37°C to digest unencapsidated nucleic acids. Nuclease treated concentrates were extracted using an SDS, KOAc, phenol-chloroform extraction and resuspended in EB Buffer (Qiagen, Buffer EB - Elution Buffer, cat# 19086) for storage at 20°C.

DNA Shearing. Phage genomic DNA was sheared by sonication in preparation for genome library preparation. DNA concentrations of extracts were determined using Picogreen (Invitrogen, Quant-iT PicoGreen dsDNA Reagent and Kits, cat# P7589) in a 96-well format and samples brought to 5µg in a 100µL final volume of PCR-grade water diluent for sonication. Samples were sonicated in batches of 6 for 6 cycles of 5 minutes each, at an interval of 30 seconds on/off on the Low Intensity setting of the Bioruptor to enrich for a fragment size of ~300bp (K2.XIV.28).

DNA Library Preparation. Illumina constructs were prepared from sheared DNA as follows: end repair of sheared DNA, 0.72x/0.21x dSPRI size selection to enrich for ~300bp sized fragments, ligation of Illumina adapters and unique pairs of forward and reverse barcodes for each sample, SPRI clean up, nick translation, and final SPRI clean up (Rodrigue et al., 2010). Constructs were enriched by PCR using PE primers following qPCR-based normalization of template concentrations. Enrichment PCRs were prepared in octuplicate 25µL volumes, with the recipe: 1µL illumina construct template, 5µL 5x Phusion polymerase buffer, 0.5µL 10mM dNTPs, 0.25µL 40µM IGA-PCR-PE-F primer, 0.25µL 40µM IGA-PCR-PE-R primer, 0.25µL Phusion polymerase, 17.75µL PCR-grade HOH. PCR thermocycling conditions were as follows: initial denaturation at 98°C for 20 sec; batch dependent number of cycles of 98°C for 15 sec, 60°C for 20 sec, 72°C for 20 sec; final annealing at 72°C for 5 min; hold at 10°C. For each sample 8 replicate enrichment PCR reactions were pooled and purified by 0.8x SPRI clean up. Each sample was then checked by Bioanalyzer (2100 expert High Sensitivity DNA Assay) to confirm the presence of a unimodal distribution of fragments with a peak between 350-500bp.

DNA Sequencing. Sequencing of phage genomes was distributed over 4 paired-end sequencing runs as follows: HiSeq library of 18 samples pooled with 18 external samples, 3 MiSeq libraries each containing ~100 multiplexed phage genomes.

CLC Genome Assembly. Raw paired-end Illumina reads were imported and demultiplexed using

CLC Genome Workbench. Duplicate PCR reads were removed using, reads were trimmed (settings: trim based on quality score = 0.01, ambiguous nucleotides = 0, remaining trim parameters set to default).

Phage Protein & Genome Clustering. Phage ORFs were called using Prodigal (Hyatt et al., 2010). Proteins were clustered using OrthoMCL (Fischer et al., 2011; Li et al., 2003) with a requirement for 75% coverage of the longer sequence (using python script written by Libusha Kelly) and an e-value cut-off of 10^{-5} . Phage genomes were clustered in Cytoscape using the FT algorithm of the ClustnSee plug-in on the basis of shared protein clusters (Shannon et al., 2003; Spinelli et al., 2013).

Electron Microscopy. All Nahant Collection lysates visualized by electron microscopy were first concentrated using a 30kD centrifugal filtration device (Millipore, Amicon Ultra Centrifugal Filters, Ultracel 30K, UFC903024). All samples were negative stained with 1-2% acidic uranyl acetate for 20-60 seconds on carbon coated ionized grids prior to viewing by transmission electron microscopy (all electron microscopy and preparation conducted by Nicki Watson [Fig. 5 in Ch. 4], except for sample 1.074.O [Panel 7 of Fig. 5 in Ch. 4], which was prepared and viewed under the guidance of Cammie Haase-Pettingell).

Additional people involved: *Optimization of Illumina library preparation and phage genome sequencing:* Radhey Shyam Sharma, Simon Labrie, Andre Mroz, Michael Cutler; *Scripts & Computational Assistance:* Libusha Kelly, Otto Cordero, Phil Arevalo, Dave VanInsberghe.

Section V. Host Range Assay

Host Range Assay Agar Overlay. A host-range matrix assay for cross infectivity was conducted to determine the potential for each of the ~300 Nahant Collection phages to infect the hosts of any of the other phages in the collection. The assay was conducted by stamping cell-free virus lysate onto host agar over-lay lawns and observing for changes in lawn morphology proximal to each stamp. Viruses were applied to host lawns using a 96-well blotter (BelArt, Bel-blotter 96-tip replicator, cat # 378760002) that was set into a microtiter plate containing arrayed phage lysate and then transferred to the surface of the host lawn and allowed to remain in contact for several minutes. Each 96-stamp contained 3 replicates of each phage lysate, distributed across three panels (columns 1-4, 5-8, 9-12) each with a unique array of the 32 samples. 96-well blotters were

microwave steam sterilized (Tommee Tippee, Closer to Nature Microwave Steam Sterilizer) in batches for continuing re-use during plating sessions. Bacterial strains were prepared for the infection assay by inoculating 1mL volumes of liquid media (2216 Marine Broth) in 2mL 96-well culture blocks directly from glycerol stocks and shaking them at RT for ~48 hours. Agar overlays were prepared by transferring ~250uL aliquots of host culture to bottom agar plates (2216 Marine Broth, 1% Bacto Agar, 5% glycerol) and adding 7mL of molten ~52C top agar (2216 Marine Broth, 0.4% Bacto Agar, 5% glycerol). Phages were prepared by distributing lysates into a 2mL 96-well culture block in panels as described above, aliquots of <200uL were then transferred into shallow microtiter plates so that the blotter could be set to rest in the phage lysates to collect phage lysate by capillary action. Host lawns were stamped with phage lysates within 5-6 hours of plating.

Host Range Assay Scoring. Agar overlays were assessed for changes in lawn morphology associated with phage treatment and scored blind with respect to phage identity and arrangement of replicates. Plates were scored for the presence of interactions on days 1, 2, 3, 7, 14, 21, and 30, and the outer bands of the interaction zones were marked with a different color for each time point. After 30 days the interactions for each strain were recorded and the approximate diameter for each interaction at each time point was recorded - either from the true diameter or from a doubling of the radius where necessary. During recording of the interactions for each plate an additional qualitative measure of confidence in the projected positive or negative call of the interaction was made. For example, where 2 of 3 replicates were positive for a phage on a lawn with no other positive interactions such an interaction would be called by the qualitative measure as “real”; alternatively, where 2 of 3 replicates were positive for a phage on a lawn containing several other positive interactions the qualitative measure might call these replicates “contam” if they were high-titer interactions and occurred in close proximity to other positive interactions.

Host Range Assay Data Entry. It is noted that for assays of this scale the time required for data entry can be as considerable as the assay itself. Due to the need to enter large numbers of datapoints for each plate (for example: 3 positive reaction measured at 7 time points = 21 data points per plate, with 15 positive reactions at 7 time points = 105 data points; this does not include additional observations regarding for example the quality of the host lawn) and large number of total plates, this amounts to thousands of single data point entries recorded manually.

It is thus highly recommended that future users of this method seek out automated methods to capture images and automate data recovery. In the case of this study, significant effort was made to minimize potential for error using specially configured Excel spreadsheets that reduced data entry to simply entering the diameter into the corresponding position in a spreadsheet array for each time point. By providing, independently, the array with position information for each of the randomized replicate IDs, and automating the re-array by lookup functions, the spreadsheet automatically generated a final output consolidating diameters by time point for all host-phage combinations per plate. This offered several stages for quality control and, despite the recommendation to consider alternate approaches in the future for work at these scales, there is high confidence in the accuracy of recording. As indicated, notes were also made regarding the quality of host lawns and observations of potential autoinduction. The full host range data set will be available in (for data location see Table 2).

REFEERENCES

- Edgar, R.C., 2004a. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113.
- Edgar, R.C., 2004b. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 32, 1792-1797.
- Fischer, S., Brunk, B.P., Chen, F., Gao, X., Harb, O.S., Iodice, J.B., Shanmugam, D., Roos, D.S., Stoeckert, C.J., 2011. Using OrthoMCL to Assign Proteins to OrthoMCL-DB Groups or to Cluster Proteomes Into New Ortholog Groups. *Current Protocols in Bioinformatics*, 6.12. 11-16.12. 19.
- Hunt, D.E., David, L.D., Gevers, D., Preheim, S.P., Alm, E.J., Polz, M.F., 2008. Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* 320, 1081-1085.
- Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 119.
- John, S.G., Mendez, C.B., Deng, L., Poulos, B., Kauffman, A.K., Kern, S., Brum, J., Polz, M.F., Boyle, E.A., Sullivan, M.B., 2011. A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environ Microbiol Rep* 3, 195-202.
- Karam, H.N., 2012. Seawater circulation in coastal aquifers : processes and impacts. *Massachusetts Institute of Technology*, 130-143.
- Letunic, I., Bork, P., 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127-128.
- Letunic, I., Bork, P., 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic acids research* 39, W475-W478.
- Li, L., Stoeckert, C.J., Roos, D.S., 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome research* 13, 2178-2189.
- Preheim, S.P., Timberlake, S., Polz, M.F., 2011. Merging taxonomy with ecological population prediction in a case study of Vibrionaceae. *Appl Environ Microbiol* 77, 7195-7206.
- Price, M.N., Dehal, P.S., Arkin, A.P., 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular biology and evolution* 26, 1641-1650.
- Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5, e9490.
- Rodrigue, S., Materna, A.C., Timberlake, S.C., Blackburn, M.C., Malmstrom, R.R., Alm, E.J., Chisholm, S.W., 2010. Unlocking short read sequencing for metagenomics. *PLoS One* 5, e11840.
- Santos, S.B., Carvalho, C.M., Sillankorva, S., Nicolau, A., Ferreira, E.C., Azeredo, J., 2009. The use of antibiotics to improve phage detection and enumeration by the double-layer agar technique. *BMC microbiology* 9, 148-148.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13, 2498-2504.
- Spinelli, L., Gambette, P., Chapple, C.E., Robisson, B., Baudot, A., Garreta, H., Tichit, L., Guénoche, A., Brun, C., 2013. Clust&See: A Cytoscape plugin for the identification, visualization and manipulation of network clusters. *BioSystems* 113, 91-95.
- Szabo, G., Preheim, S.P., Kauffman, K.M., David, L.A., Shapiro, J., Alm, E.J., Polz, M.F., 2012. Reproducibility of Vibrionaceae population structure in coastal bacterioplankton. *ISME J*.

FIGURES AND TABLES



Figure 1. Nahant Time Series sampling site. Red arrow indicates Canoe Cove on Nahant (42.41975, -70.907114), the site at which all samples were collected. Map and coordinates by Google Maps (April 2014).



Station 3



Station 2



Station 1

Figure 2. Three sampling “stations” at Canoe Cove.

Nahant 2010									
Field Personnel	Lab Personnel	Field - Time, Arrival:	Field - Time, Departure:	Weekday	Month	Day	Year	Julian Day	Other Notes
KK/	KK/	10:25		SATURDAY	OCTOBER	23	2010		296!
		Lab - Time, Arrival:	Lab - Time, Departure:	Air Temp	HOH Temp	Salinity	Weather & Tide Notes		
				8.7 A	11.2 D				
				8.9 D	11.2 A				
PICTURE:									

Station	Sample #	Confirm #	Sample Description	2mL Formalin Added?	On Ice?	Notes	Time stored	Storage Location
Station 0	1	/	Counts - A (10mL) + 2mL 0.02 formalin in 125mL bottle	/	/	DIE		Basement / New 40L Bottom Shelf
	2	/	Counts - B (10mL) + 2mL 0.02 formalin in 125mL bottle	/	/	~ 1315-1420?		
	3	/	Counts - C (10mL) + 2mL 0.02 formalin in 125mL bottle	/	/			
Station 0	4	/	DOC - A (10mL) in screw cap culture tube	/	/			40L Floor / Work in 40L cold room / Right side (in dark box)
	5	/	DOC - B (10mL) in screw cap culture tube	/	/			
	6	/	DOC - C (10mL) in screw cap culture tube	/	/			
Station 1	7	/	Chemicals - A (150mL) through ashed GF/F into acid washed bottle	/	/			40L / 2nd -30C / Bottom Shelf
	8	/	Chemicals - B (150mL) through ashed GF/F into acid washed bottle	/	/			
	9	/	Chemicals - C (150mL) through ashed GF/F into acid washed bottle	/	/			
Station 1	10	/	Chlorophyll - A (500mL) onto GF/C filter	/	/			Storage Location (Filter in 4L Fall)
	11	/	Chlorophyll - B (500mL) onto GF/C filter	/	/			
	12	/	Chlorophyll - C (500mL) onto GF/C filter	/	/			
Station 1	13	/	24.7mL DNA - 63mL Fraction - A (100L in cask) → Grind and set to exactly 30mL, syringe filter 10mL through (retrieved)	/	/			Storage Location (Filter in microtub)
	14	/	28mL DNA - 63mL Fraction - B (100L in cask) → Grind and set to exactly 30mL, syringe filter 10mL through (survive)	/	/			
	15	/	35mL DNA - 63mL Fraction - C (100L in cask) → Grind and set to exactly 30mL, syringe filter 10mL onto 0.2um PC filter	/	/			
Station 1	16	/	Glycerol - 63mL Fraction - A (100L in cask) → Grind and set exactly 30mL, filter 20mL onto 0.2um PC filter	/	/			Storage Location (40L tube)
	17	/	Glycerol - 63mL Fraction - B (100L in cask) → Grind and set exactly 30mL, filter 20mL onto 0.2um PC filter	/	/			
	18	/	Glycerol - 63mL Fraction - C (100L in cask) → Grind and set exactly 30mL, filter 20mL onto 0.2um PC filter	/	/			

Dumped an extra ~ 1ul out for pix of zoops.

Figure 3 (panel 1 of 4). Example daily datasheet from Nahant Time Series. For each Sample ID there is a brief description of protocol used.

Collected	Confirm #	Sample #	Sample Description	Fiber Log	Volume	Notes	Time stored	Storage Location (Filter in microtube)
Station 2	/	19	DNA - 3uM fraction - A (4) filtered onto a 3uM PC	WV 1000 2619			1445	418 / 2nd -20C / Nahant 2010 DNA filters Box
	/	20	DNA - 3uM fraction - B (4) filtered onto a 3uM PC				1447	
	/	21	DNA - 3uM fraction - C (4) filtered onto a 3uM PC				1447	
Station 2	/	22	DNA - 3uM fraction - A (4) filtered onto a 3uM PC	SKN 1000 2619			1453	418 / 2nd -20C / Nahant 2010 DNA filters Box
	/	23	DNA - 3uM fraction - B (4) filtered onto a 3uM PC				1453	
	/	24	DNA - 3uM fraction - C (4) filtered onto a 3uM PC				1453	
Station 2	/	25	DNA - 0.2uM fraction - A (4) filtered through a 3uM PC	WV 1000 2619			1500	418 / 2nd -20C / Nahant 2010 DNA filters Box
	/	26	DNA - 0.2uM fraction - B (4) filtered through a 3uM PC				1500	
	/	27	DNA - 0.2uM fraction - C (4) filtered through a 3uM PC				1500	
Station 2	/	28	Glycerol - 3uM fraction - A (4) filtered onto a 3uM PC	WV 1000 2619			1816	Basement / New 40 2nd from Bottom Shelf / Nahant 2010 Community Glycerol Box
	/	29	Glycerol - 3uM fraction - B (4) filtered onto a 3uM PC				1816	
	/	30	Glycerol - 3uM fraction - C (4) filtered onto a 3uM PC				1816	
Station 2	/	31	Glycerol - 3uM fraction - A (4) filtered onto a 3uM PC	WV 1000 2619			1825	Basement / New 40 2nd from Bottom Shelf / Nahant 2010 Community Glycerol Box
	/	32	Glycerol - 3uM fraction - B (4) filtered onto a 3uM PC				1825	
	/	33	Glycerol - 3uM fraction - C (4) filtered onto a 3uM PC				1825	
Station 2	/	34	Glycerol - 0.2uM fraction - A (4) filtered onto a 0.2 PC	WV 1000 2619			1857	Basement / New 40 2nd from Bottom Shelf / Nahant 2010 Community Glycerol Box
	/	35	Glycerol - 0.2uM fraction - B (4) filtered onto a 0.2 PC				1857	
	/	36	Glycerol - 0.2uM fraction - C (4) filtered onto a 0.2 PC				1857	
Station 3	Collected	37	DNA - Unfractionated HOH - A (4) filtered through a 3uM PC				1345	418 / 2nd -20C / Nahant 2010 DNA filters Box
	/	38	DNA - Unfractionated HOH - B (4) filtered through a 3uM PC				1345	
	/	39	DNA - Unfractionated HOH - C (4) filtered through a 3uM PC				1345	
Station 3	Collected	40	Phage - A, from 0.2 uM PC, add 400uL FACS solution, let sit 1hr, filter onto 0.2uM PC, remove filter to 7mL, 85 uL, resuspended in 4mL, on-ice				1345	418 / 2nd -20C / Nahant 2010 DNA filters Box
	/	41	Phage - B, from 0.2 uM PC, add 400uL FACS solution, let sit 1hr, filter onto 0.2uM PC, remove filter to 7mL, 85 uL, resuspended in 4mL, on-ice				1345	
	/	42	Phage - C, from 0.2 uM PC, add 400uL FACS solution, let sit 1hr, filter onto 0.2uM PC, remove filter to 7mL, 85 uL, resuspended in 4mL, on-ice				1345	

Figure 3 (panel 2 of 4). Example daily datasheet from Nahant Time Series. For each Sample ID there is a brief description of protocol used.

Comments Page for Julian Day:

296.

I can see many long hairs / rods in HOT for start	bucket !!! amazing, big block
water clear / light green?	
some little mangrove debris @ 2	divers @ end (not in hot)
1 or 2 compressed swimmers	going to "chirney"
landward beach pickup	
light orange haze	crows eating.
hazy, no clouds	
couple gulls	lots of green/bleaching toward
lots of ducks @ mouth	dicholap??
high tide incoming	

Daily Supply Check List

- Car Keys
- MIT ID (Garage) & License
- Individual liabilities sheets
- Cash for tolls (\$\$)
- Tail Socks
- Personal neoprene socks
- Clipboard
- Tape
- Folder with data sheet + directions, pen
- 1 Digital Thermometer
- 2 Alcohol Thermometer
- Refractometer
- Bucket
- Squirt Bottle
- 63uM net
- 2 pairs of waders
- camera
- 3 125 ml bottles (counts)
- 1+ falcon tubes
- Formalin in 15ml falcon
- Pipette
- Pipette tips
- Cooler
- with ice packs
- 1+ 10 ml syringes
- 1+ Acrodisk 0.2 filters
- 3+ screwcap vials (DOC)
- Baggy to put above vials in
- 3 250 ml acid-washed bottles (5 if doing controls)
- 3 1 L bottles (Chlorophyll)
- 3 oakridge tubes
- 3+ 3L bottles
- 3+ 4L bottles
- white rack for thermometers
- Dry ice added to Ethanol?
- 3x16 L
- expressies
- dump pile

Figure 3(panel 3 of 4). Example daily datasheet from Nahant Time Series. For each Sample ID there is a brief description of protocol used.

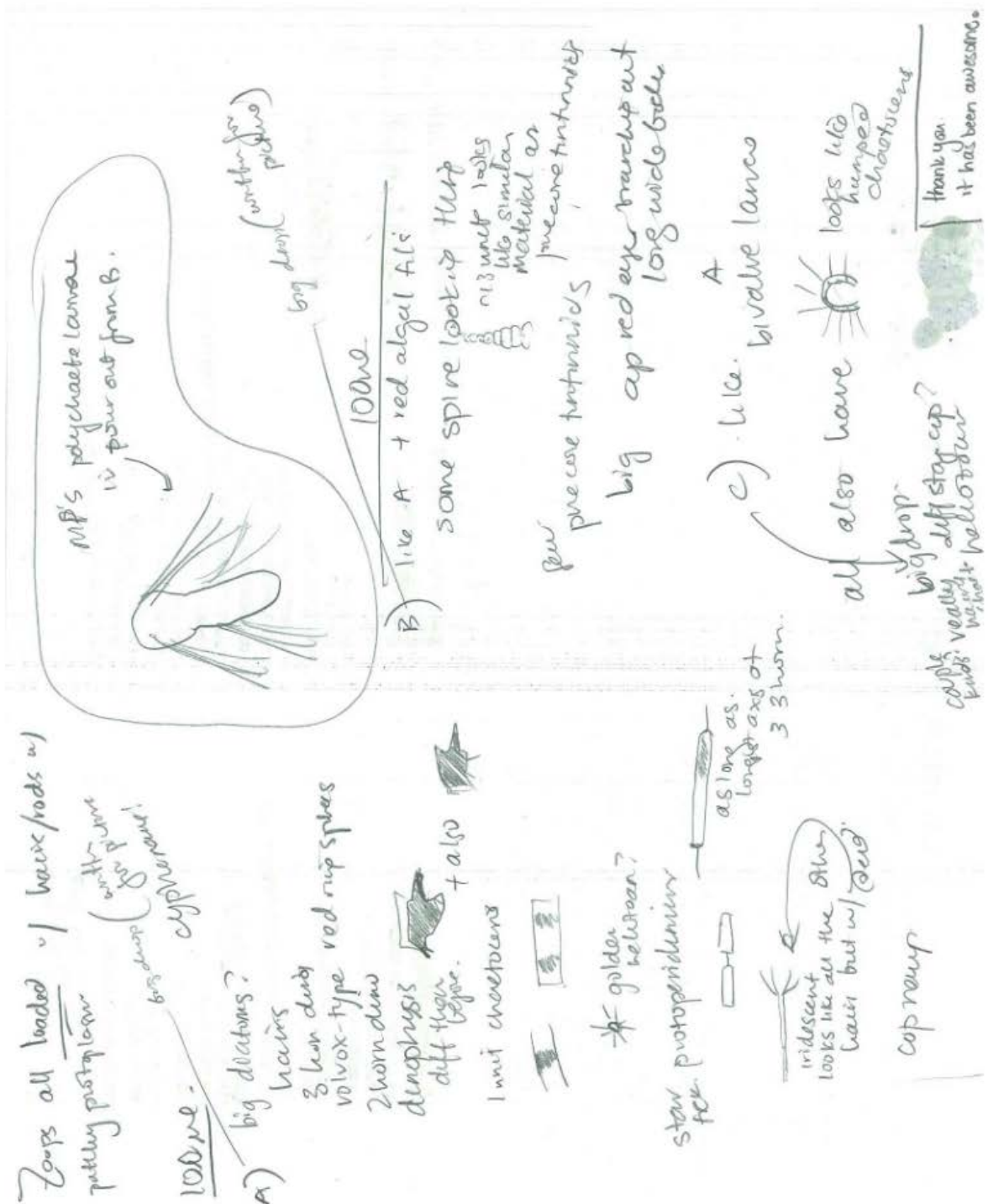


Figure 3 (panel 4 of 4). Example daily datasheet from Nahant Time Series. For each Sample ID there is a brief description of protocol used.

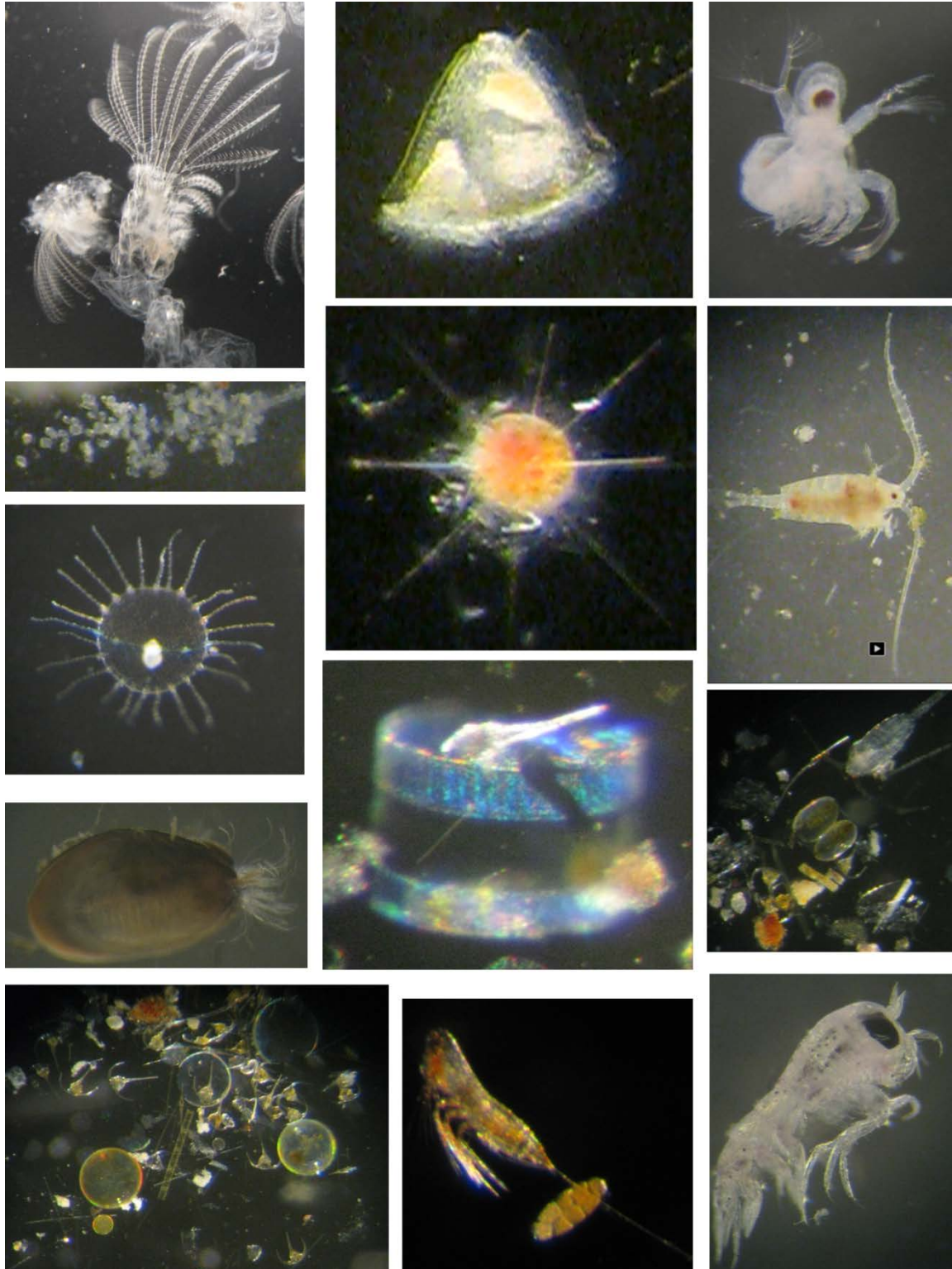


Figure 4. Sample of images available in Nahant Collection microscopy photo gallery. Images not to scale.

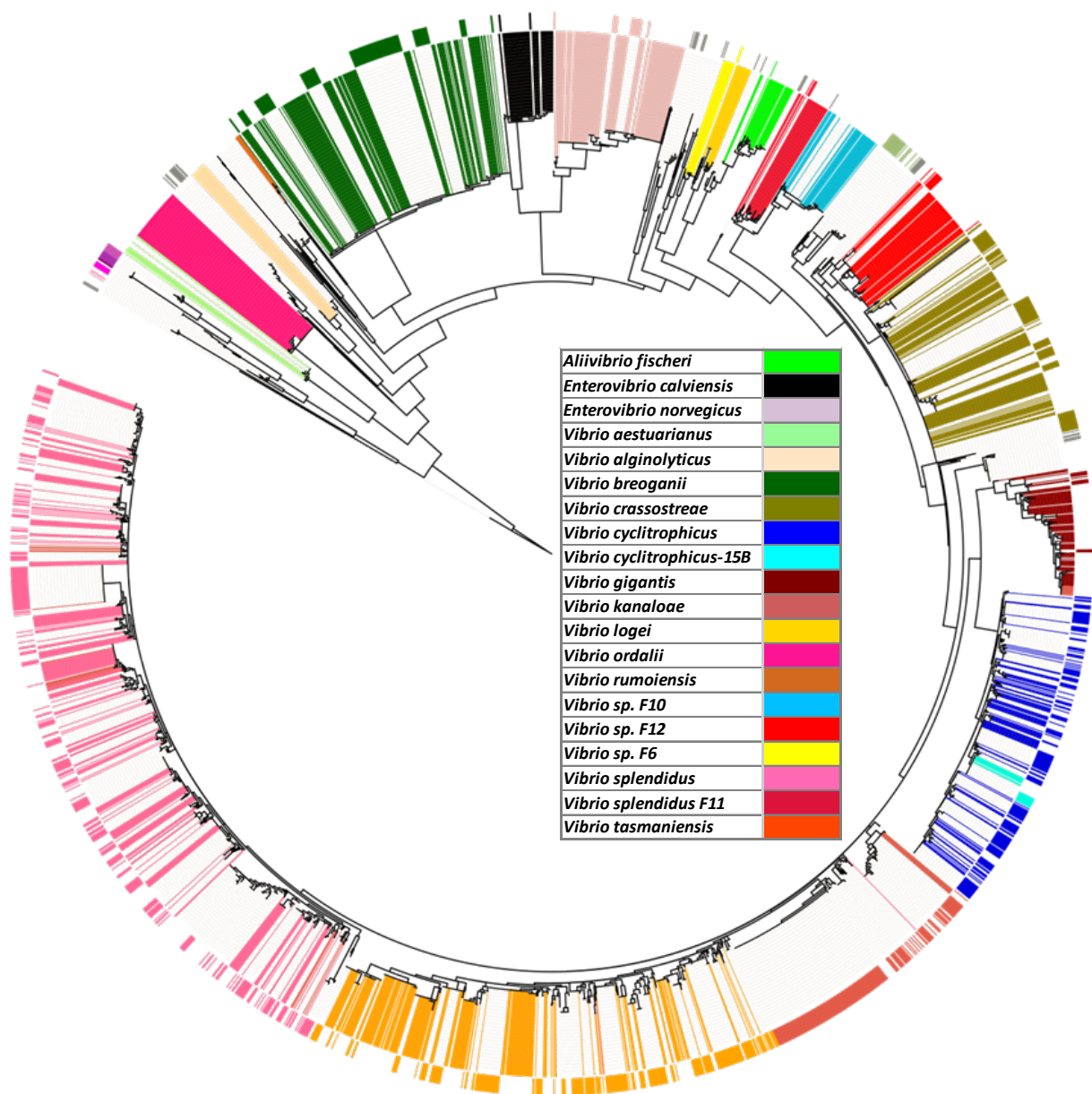


Figure 5 Phylogenetic tree of Nahant Collection isolates in relation to previous isolates, based on partial *Hsp60* sequence. Outermost color ring indicates Nahant strains, all other strains derive from historical collections (Hunt et al., 2008; Szabo et al., 2012). Color mappings indicate population assignment.

Table 1. Locations of data associated with the Nahant Collection

Host:	polzserver.mit.edu
Main Directory:	/Users/data_access/PolzData/NahantCollection/
Sub Directory	Description
/SAMPLING/01.Field_data_sheets/	Raw scanned datasheets.
/SAMPLING/02.Datasets/	Datasets largely transcribed from raw datasheets.
/SAMPLING/03.NOAA/	NOAA buoy station datasets corresponding to sampling period.
/SAMPLING/04.Pictures/	Field and microscopy images
/GENOMES/Hosts/	Host genomes and derived data.
/GENOMES/Phages/	Phage genomes and derived data.
/PHAGE/01.Master_Interactive	Interactive spreadsheet integrating multiple layers of data about phages and intended to be used as a database when looking for phages with particular characteristics relating to host range, co-occurrence, host population, genome size, genome cluster, etc... Filterable for searching for specific subsets of interest.
/PHAGE/02.LSP/	Included files: Summary catalog of archived and described plaques; Raw LSP data
/PHAGE/03.HRX/	Included files: Example semi-automated data entry sheet for use with large scale host range assays; Raw host range data

Table 2. Nomenclature for Nahant Collection bacterial isolates.

	Replicate 1	Replicate 2	Replicate 3
63uM	54 ¹	55	56
	Replicate A	Replicate B	Replicate C
5.0uM	51	52	53
1.0uM	48	49	50
0.2uM	45	46	47

¹Glycerol stocks of strains are stored in plates with the nomenclature 10N.xxx.yy, where xxx is the ordinal day of collection and yy is the above indicated plate number which can be used to determine the size fraction as well which other size fraction samples derive from the same physical replicate. Samples for the 63uM isolations were independent of those for the smaller size fractions.

Table 3 (panel 1 of 7). Catalog of Nahant Collection phages and isolation hosts¹.

Phage ID	Strain ID	PFU 15mL ⁻¹	Hsp60-based population assignment	Phage Genome Cluster (Thesis)	Genome Length
1.001.	10N.261.46.E5	AI-TMTC	<i>Vibrio tasmaniensis</i>		
1.002.	10N.261.52.F1	AI-TMTC	<i>Vibrio tasmaniensis</i>		
1.003.O.	10N.286.48.A2	127	<i>Vibrio tasmaniensis</i>	1	41,891
1.004.O.	10N.261.54.A2	81	<i>Vibrio tasmaniensis</i>	14	45,496
1.005.O.	10N.286.48.F2	69	<i>Vibrio splendidus</i>	7	51,088
1.006.	10N.286.55.B4	65	<i>Vibrio cyclitrophicus</i>		
1.007.O.	10N.261.55.F9	56	<i>Vibrio cyclitrophicus</i>	1	49,244
1.008.O.	10N.286.54.E5	49	<i>Vibrio cyclitrophicus</i>	15	10,579
1.009.O.	10N.261.51.C9	44	<i>Vibrio tasmaniensis</i>	7	44,428
1.011.O.	10N.286.49.B11	37	<i>Vibrio cyclitrophicus</i>	15	10,579
1.012.O.	10N.261.48.C12	19	<i>Vibrio tasmaniensis</i>	4	59,851
1.013.O.	10N.286.54.F9	19	<i>Vibrio cyclitrophicus</i>	9	44,457
1.015.O.	10N.222.51.E5	18	<i>Vibrio breoganii</i>	8	42,586
1.016.O.	10N.286.46.A11	18	<i>Vibrio splendidus</i>	1	48,000
1.017.O.	10N.286.55.C11	17	<i>Vibrio sp. F12</i>	7	51,290
1.020.O.	10N.222.48.A2	1	<i>Vibrio kanaloae</i>	15	10,578
1.021.A.	10N.222.51.F9	1	<i>Vibrio tasmaniensis</i>	11	44,174
1.021.B.	10N.222.51.F9	1	<i>Vibrio tasmaniensis</i>	11	44,534
1.021.C.	10N.222.51.F9	1	<i>Vibrio tasmaniensis</i>	11	44,309
1.022.O.	10N.286.45.A10	1	<i>Vibrio splendidus</i>	4	62,440
1.023.O.	10N.222.51.B4	1	<i>Vibrio splendidus</i>	4	59,872
1.024.O.	10N.261.45.F8	8	<i>Vibrio splendidus</i>	4	58,831
1.025.O.	10N.222.46.B6	2	<i>Vibrio splendidus</i>	5	75,797
1.026.O.	10N.222.49.C7	6	<i>Vibrio splendidus</i>	5	75,818
1.027.O.	10N.286.54.B8	2	<i>Vibrio splendidus</i>	4	59,297
1.028.O.	10N.286.45.B6	1	<i>Vibrio sp. 10N-1</i>	17	31,801
1.029.O.	10N.261.55.A7	14	unmapped	1	46,513
1.030.O.	10N.222.55.F9	2	<i>Vibrio crassostreae</i>	7	45,718
1.031.O.	10N.261.46.F8	1	<i>Vibrio tasmaniensis</i>	3	152,971
1.032.O.	10N.261.54.F5	2	<i>Vibrio kanaloae</i>	4	60,399
1.033.O.	10N.222.49.B8	8	<i>Vibrio kanaloae</i>	4	61,206
1.034.O.	10N.261.46.B7	4	<i>Vibrio breoganii</i>	8	44,380
1.034.X.	10N.261.46.B7	4	<i>Vibrio breoganii</i>	8	44,380
1.035.	10N.286.49.C5	AI-176	<i>Vibrio tasmaniensis</i>		
1.036.O.	10N.286.45.C3	61	<i>Vibrio tasmaniensis</i>	1	40,172
1.037.O.	10N.261.52.F7	51	<i>Vibrio tasmaniensis</i>	14	43,111
1.038.O.	10N.286.51.C2	50	<i>Vibrio tasmaniensis</i>	1	49,567
1.038.O.	10N.286.51.C2	50	<i>Vibrio tasmaniensis</i>	1	49,567
1.039.O.	10N.286.55.A2	48	<i>Vibrio tasmaniensis</i>	1	40,972
1.040.O.	10N.286.45.B9	43	<i>Vibrio tasmaniensis</i>	15	10,587
1.041.	10N.261.55.E8	42	<i>Vibrio tasmaniensis</i>		
1.042.O.	10N.286.45.B8	28	<i>Vibrio tasmaniensis</i>	1	49,736
1.043.O.	10N.261.52.C7	18	<i>Vibrio tasmaniensis</i>	15	10,272
1.044.O.	10N.261.51.B8	10	<i>Vibrio tasmaniensis</i>	15	10,272
2.044.O.	10N.261.51.B8	10	<i>Vibrio tasmaniensis</i>	14	45,132
1.046.O.	10N.286.52.E3	2	<i>Vibrio splendidus</i>	4	62,503
1.047.O.	10N.286.55.F2	2	<i>Vibrio cyclitrophicus</i>	1	46,052

¹Includes column for number of plaque forming units (PFUs) observed on host strain during Large Scale Plating; notation of AI indicates strain is known autoinducer; TMTC indicates “too many to count”. Genome cluster as per Chapter 4 of this thesis.

Table 3 (panel 2 of 7). Catalog of Nahant Collection phages and isolation hosts

Phage ID	Strain ID	PFU 15mL ⁻¹	Hsp60-based population assignment	Phage Genome Cluster (Thesis)	Genome Length
1.048.O.	10N.286.46.A10	13	<i>Vibrio tasmaniensis</i>	15	10,460
1.049.O.	10N.286.54.B5	439	<i>Shewanella</i> sp. 9ZD10-like G2	12	45,347
1.050.O.	10N.286.48.A6	93	<i>Shewanella</i> sp. 9ZD10-like G1	12	45,233
1.051.	10N.222.52.A11	AI-90	<i>Vibrio tasmaniensis</i>		
1.052.A.	10N.286.46.C3	74	<i>Enterovibrio norvegicus</i>	1	42,889
1.054.O.	10N.261.52.A1	70	<i>Vibrio</i> sp. 10N-1	1	41,729
1.055.O.	10N.286.55.E9	64	<i>Enterovibrio norvegicus</i>	12	47,406
1.056.O.	10N.261.48.C11	57	<i>Vibrio tasmaniensis</i>	14	47,054
1.057.O.	10N.261.46.B12	49	<i>Vibrio tasmaniensis</i>	15	10,273
2.058.O.	10N.286.46.B8	41	<i>Vibrio tasmaniensis</i>	3	101,637
3.058.O.	10N.286.46.B8	41	<i>Vibrio tasmaniensis</i>	3	101,642
1.059.A.	10N.286.48.A9	39	<i>Vibrio tasmaniensis</i>	1	41,540
1.060.A.	10N.261.48.B5	35	<i>Vibrio splendidus</i> F11	1	41,087
1.061.O.	10N.286.55.C2	35	<i>Vibrio tasmaniensis</i>	1	41,667
1.062.O.	10N.286.55.C3	35	<i>Vibrio cyclitrophicus</i>	15	10,579
1.063.O.	10N.261.45.C7	32	<i>Vibrio tasmaniensis</i>	2	128,645
1.064.O.	10N.261.52.E5	27	<i>Vibrio crassostreae</i>	6	36,584
1.065.	10N.261.45.B10	24	<i>Vibrio tasmaniensis</i>		
1.066.O.	10N.286.46.E8	24	<i>Vibrio tasmaniensis</i>	14	47,053
1.067.O.	10N.261.52.C9	23	<i>Vibrio tasmaniensis</i>	1	40,955
1.068.O.	10N.261.51.F8	22	<i>Vibrio tasmaniensis</i>	14	46,975
1.069.O.	10N.286.49.F11	21	<i>Vibrio tasmaniensis</i>	15	10,644
1.070.O.	10N.261.45.B2	19	<i>Enterovibrio norvegicus</i>	8	44,087
1.071.A.	10N.286.46.A12	19	<i>Vibrio splendidus</i>	1	48,114
1.071.B.	10N.286.46.A12	19	<i>Vibrio splendidus</i>	1	48,114
1.071.C.	10N.286.46.A12	19	<i>Vibrio splendidus</i>	1	48,114
1.071.D.	10N.286.46.A12	19	<i>Vibrio splendidus</i>	1	48,051
1.072.O.	10N.286.48.A12	18	unmapped	14	43,788
1.073.O.	10N.286.55.C1	17	<i>Vibrio tasmaniensis</i>	1	41,507
1.074.O.	10N.222.49.B7	15	<i>Vibrio breoganii</i>	6	36,922
1.075.O.	10N.286.55.B10	15	<i>Vibrio</i> sp. F12	7	51,228
1.076.O.	10N.286.51.B7	14	<i>Shewanella</i> sp. 9ZD10-like G2	9	47,914
1.077.O.	10N.261.45.A10	13	<i>Enterovibrio norvegicus</i>	8	44,047
1.078.O.	10N.261.51.C4	12	<i>Vibrio crassostreae</i>	1	42,417
1.079.O.	10N.286.45.E9	12	<i>Vibrio tasmaniensis</i>	1	42,312
1.080.O.	10N.286.48.A4	12	<i>Vibrio tasmaniensis</i>	15	10,046
1.081.O.	10N.286.52.C2	11	<i>Shewanella</i> sp. 9ZD10-like G1	3	239,318
1.082.O.	10N.261.49.E4	10	<i>Vibrio breoganii</i>	6	35,683
1.083.O.	10N.286.52.B9	10	<i>Shewanella</i> sp. 9ZD10-like G1	12	45,364
1.084.O.	10N.261.49.F5	9	<i>Enterovibrio norvegicus</i>	2	141,906
1.085.O.	10N.222.51.E3	8	<i>Vibrio breoganii</i>	6	37,820
1.086.O.	10N.222.51.F8	8	<i>Vibrio tasmaniensis</i>	7	50,835
1.087.A.	10N.261.45.F7		unmapped	7	45,674
1.087.B.	10N.261.45.F7		unmapped	7	45,674
1.088.O.	10N.261.46.A1	8	<i>Vibrio kanaloae</i>	4	60,385
1.089.O.	10N.261.51.F9	8	<i>Vibrio tasmaniensis</i>	4	59,851
1.090.B.	10N.286.48.F1	8	<i>Vibrio tasmaniensis</i>	1	41,806

Table 3 (panel 3 of 7). Catalog of Nahant Collection phages and isolation hosts.

Phage ID	Strain ID	PFU 15mL ⁻¹	Hsp60-based population assignment	Phage Genome Cluster (Thesis)	Genome Length
1.091.O.	10N.286.52.B12	8	<i>Vibrio splendidus</i>	1	43,134
2.092.O.	10N.286.52.B7	8	<i>Vibrio tasmaniensis</i>	15	10,580
1.093.O.	10N.286.55.E10	8	<i>Vibrio sp. F12</i>	7	51,290
1.094.O.	10N.286.55.E12	8	<i>Vibrio sp. F12</i>	7	51,290
1.095.O.	10N.286.46.E10	7	<i>Vibrio sp. F12</i>	15	10,436
2.095.A.	10N.286.46.E10	7	<i>Vibrio sp. F12</i>	1	44,649
2.095.B.	10N.286.46.E10	7	<i>Vibrio sp. F12</i>	1	44,649
2.096.O.	10N.286.48.B5	7	<i>Shewanella sp. 9ZD10-like G1</i>	12	44,991
1.097.O.	10N.286.49.B3	7	unmapped	5	76,918
1.098.O.	10N.286.51.B9	7	<i>Vibrio tasmaniensis</i>	4	59,851
1.100.O.	10N.261.45.C3	6	<i>Vibrio tasmaniensis</i>	7	51,268
1.101.O.	10N.261.45.C6	6	<i>Enterovibrio norvegicus</i>	2	130,187
1.102.O.	10N.261.45.E3	6	<i>Vibrio tasmaniensis</i>	15	10,447
1.103.O.	10N.261.52.F2	6	<i>Vibrio kanaloae</i>	4	61,759
1.104.O.	10N.286.49.A12	6	<i>Vibrio tasmaniensis</i>	4	60,990
1.105.O.	10N.286.49.B4	6	<i>Vibrio cyclitrophicus</i>	1	49,043
1.106.O.	10N.286.51.F7	6	<i>Vibrio cyclitrophicus</i>	1	47,070
1.107.A.	10N.286.52.E10	6	<i>Vibrio tasmaniensis</i>	15	10,447
1.107.B.	10N.286.52.E10	6	<i>Vibrio tasmaniensis</i>	15	10,447
1.107.C.	10N.286.52.E10	6	<i>Vibrio tasmaniensis</i>	15	10,447
1.108.O.	10N.222.51.A4	5	unmapped	6	36,584
1.109.O.	10N.261.46.C3	5	<i>Vibrio breoganii</i>	6	37,347
1.110.O.	10N.261.52.C1	5	<i>Vibrio breoganii</i>	6	37,556
1.111.A.	10N.286.45.E6	5	<i>Vibrio tasmaniensis</i>	18	40,209
1.111.B.	10N.286.45.E6	5	<i>Vibrio tasmaniensis</i>	18	40,344
1.112.O.	10N.286.46.B11	5	<i>Vibrio tasmaniensis</i>	1	48,149
1.113.A.	10N.286.51.E7	5	<i>Vibrio kanaloae</i>	1	44,144
1.114.O.	10N.286.55.B8	5	<i>Vibrio kanaloae</i>	1	37,451
1.115.A.	10N.222.49.B11	4	<i>Vibrio breoganii</i>	6	37,416
1.115.B.	10N.222.49.B11	4	<i>Vibrio breoganii</i>	6	37,362
1.116.O.	10N.222.52.C10	4	<i>Vibrio breoganii</i>	6	36,314
1.117.O.	10N.261.45.E8	4	<i>Vibrio kanaloae</i>	13	55,795
2.117.O.	10N.261.45.E8	4	<i>Vibrio kanaloae</i>	13	55,796
1.118.A.	10N.261.49.F6	4	<i>Vibrio tasmaniensis</i>	4	60,656
1.118.B.	10N.261.49.F6	4	<i>Vibrio tasmaniensis</i>	4	60,458
1.119.O.	10N.261.51.A9	4	<i>Vibrio splendidus</i>	1	44,527
1.121.O.	10N.286.46.C4	4	<i>Enterovibrio norvegicus</i>	2	145,527
1.122.A.	10N.286.46.F8	4	<i>Vibrio cyclitrophicus</i>	9	44,523
1.122.B.	10N.286.46.F8	4	<i>Vibrio cyclitrophicus</i>	9	44,523
1.123.O.	10N.286.48.F3	4	<i>Enterovibrio norvegicus</i>	1	46,071
1.124.O.	10N.286.49.B1	4	<i>Vibrio splendidus</i>	7	47,683
1.125.O.	10N.286.49.F5	4	<i>Vibrio splendidus</i>	15	10,578
1.126.O.	10N.286.51.A1	4	<i>Vibrio tasmaniensis</i>		
1.127.O.	10N.286.52.E12	4	<i>Vibrio splendidus</i>	1	44,915
2.130.O.	10N.222.46.C2	3	<i>Vibrio splendidus</i>	5	75,797
1.131.O.	10N.222.49.A8	3	<i>Vibrio breoganii</i>	6	37,933
1.132.O.	10N.222.49.F8	3	<i>Vibrio breoganii</i>	6	36,926

Table 3 (panel 4 of 7). Catalog of Nahant Collection phages and isolation hosts.

Phage ID	Strain ID	PFU 15mL ⁻¹	Hsp60-based population assignment	Phage Genome Cluster (Thesis)	Genome Length
1.133.O.	10N.222.51.E4	3	<i>Vibrio breoganii</i>	6	37,087
1.134.O.	10N.222.52.B8	3	<i>Vibrio breoganii</i>	6	36,379
1.135.O.	10N.222.54.B6	3	<i>Vibrio crassostreae</i>	1	41,918
1.136.O.	10N.261.45.E11	3	<i>Vibrio sp. 10N-1</i>	17	31,781
1.137.O.	10N.261.46.B5	3	<i>Vibrio splendidus</i>	1	37,958
1.138.O.	10N.261.48.A1	3	<i>Vibrio tasmaniensis</i>	11	32,510
1.139.A.	10N.261.48.C6	3	<i>Vibrio breoganii</i>	8	43,911
1.139.B.	10N.261.48.C6	3	<i>Vibrio breoganii</i>	8	43,974
1.140.	10N.261.49.B10	3	<i>Vibrio sp. F6</i>		
1.141.A.	10N.261.49.B3	3	<i>Vibrio gigantis</i>	15	10,047
1.142.O.	10N.261.49.E11	3	<i>Vibrio sp. 10N-1</i>	17	31,758
1.143.O.	10N.261.55.C8	3	<i>Vibrio splendidus</i>	1	44,527
1.144.O.	10N.286.45.B3	3	<i>Vibrio splendidus</i>	1	44,418
1.145.O.	10N.286.45.F7	3	<i>Vibrio cyclitrophicus</i>	7	51,171
1.147.O.	10N.286.49.E9	3	unmapped	6	37,360
1.148.O.	10N.286.54.A10	3	<i>Vibrio breoganii</i>	6	36,656
1.149.O.	10N.286.55.A12	3	<i>Vibrio cyclitrophicus</i>	1	48,481
1.150.O.	10N.222.46.A6	2	<i>Vibrio splendidus</i>	5	75,796
1.151.O.	10N.222.46.B1	2	<i>Vibrio splendidus</i>	1	44,307
1.152.O.	10N.222.46.E1	2	<i>Vibrio splendidus</i>	5	75,736
2.153.O.	10N.222.49.A5	2	unmapped	6	37,256
1.154.O.	10N.222.52.B12	2	unmapped	1	37,073
1.155.O.	10N.222.55.B3	2	<i>Vibrio crassostreae</i>	16	29,029
1.156.O.	10N.261.45.A6	2	<i>Vibrio sp. 10N-1</i>	17	31,680
1.157.O.	10N.261.45.B7	2	<i>Vibrio breoganii</i>	6	35,855
1.158.O.	10N.261.45.E12	2	<i>Vibrio tasmaniensis</i>	7	50,299
1.159.O.	10N.261.46.F12	2	<i>Vibrio sp. 10N-1</i>	17	31,620
2.159.A.	10N.261.46.F12	2	<i>Vibrio sp. 10N-1</i>	17	31,819
2.159.B.	10N.261.46.F12	2	<i>Vibrio sp. 10N-1</i>	17	31,679
1.160.O.	10N.261.48.B11	2	<i>Vibrio tasmaniensis</i>	4	60,734
1.161.O.	10N.261.48.C5	2	unmapped	2	140,668
1.162.O.	10N.261.48.E3	2	<i>Vibrio breoganii</i>	6	37,360
1.164.O.	10N.261.51.A7	2	<i>Vibrio crassostreae</i>	7	48,217
1.165.O.	10N.261.51.B7	2	unmapped	6	37,046
1.166.O.	10N.261.51.C7	2	<i>Vibrio kanaloae</i>	16	21,800
1.167.O.	10N.261.51.F2	2	<i>Vibrio breoganii</i>	6	36,086
1.168.O.	10N.261.52.A10	2	<i>Vibrio breoganii</i>	6	37,551
1.169.O.	10N.261.52.B1	2	<i>Vibrio breoganii</i>	5	72,307
1.170.O.	10N.261.52.C3	2	<i>Vibrio splendidus</i>	2	133,692
1.171.O.	10N.261.52.F12	2	<i>Vibrio splendidus</i>	4	60,274
1.172.O.	10N.261.52.F5	2	<i>Vibrio breoganii</i>	6	36,314
1.173.O.	10N.261.55.A11	2	<i>Vibrio cyclitrophicus</i>	1	48,063
1.174.O.	10N.261.55.A8	2	<i>Vibrio tasmaniensis</i>	7	47,102
1.175.O.	10N.261.55.B3	2	<i>Vibrio tasmaniensis</i>	7	49,066
1.176.O.	10N.261.55.F5	2	<i>Vibrio breoganii</i>	6	36,463
1.177.O.	10N.286.45.E10	2	<i>Vibrio sp. F12</i>	1	45,439
1.178.O.	10N.286.45.E12	2	<i>Vibrio cyclitrophicus</i>	1	39,934

Table 3 (panel 5 of 7). Catalog of Nahant Collection phages and isolation hosts.

Phage ID	Strain ID	PFU 15mL ⁻¹	Hsp60-based population assignment	Phage Genome Cluster (Thesis)	Genome Length
1.179.O.	10N.286.45.F12	2	<i>Vibrio splendidus</i>	11	31,498
1.181.O.	10N.286.46.C9	2	<i>Vibrio splendidus</i>	7	50,228
1.182.O.	10N.286.46.E1	2	<i>Vibrio breoganii</i>	6	36,910
1.183.O.	10N.286.48.B7	2	<i>Vibrio splendidus</i>	9	37,505
1.184.A.	10N.286.49.A5	2	<i>Vibrio cyclitrophicus</i>	9	33,277
1.184.B.	10N.286.49.A5	2	<i>Vibrio cyclitrophicus</i>	9	33,541
1.185.O.	10N.286.49.C2	2	<i>Vibrio splendidus</i>	11	43,601
1.186.O.	10N.286.49.E3	2	<i>Vibrio cyclitrophicus</i>	1	48,643
1.187.O.	10N.286.49.F1	2	<i>Vibrio kanaloae</i>	2	133,254
1.188.A.	10N.286.51.A6	2	<i>Vibrio breoganii</i>	5	72,305
1.188.B.	10N.286.51.A6	2	<i>Vibrio breoganii</i>	5	72,242
1.188.C.	10N.286.51.A6	2	<i>Vibrio breoganii</i>	5	72,305
1.189.B.	10N.286.51.B5	2	<i>Vibrio breoganii</i>	6	36,855
1.189.C.	10N.286.51.B5	2	<i>Vibrio breoganii</i>	6	36,855
1.189.O.	10N.286.51.B5	2	<i>Vibrio breoganii</i>	6	36,855
1.190.O.	10N.286.51.F12	2	<i>Vibrio kanaloae</i>	16	21,785
1.191.O.	10N.286.52.B4	2	<i>Vibrio splendidus</i>	7	48,354
1.192.O.	10N.286.52.C12	2	<i>Vibrio breoganii</i>	10	76,093
1.193.O.	10N.286.52.C6	2	<i>Vibrio splendidus</i>	2	132,560
1.194.O.	10N.286.54.B1	2	<i>Vibrio splendidus</i>	4	59,294
1.195.O.	10N.286.54.C8	2	<i>Vibrio splendidus</i>	4	59,738
1.196.O.	10N.286.54.E12	2	<i>Vibrio splendidus</i>	4	59,414
1.197.A.	10N.286.54.F2	2	<i>Vibrio cyclitrophicus</i>	1	44,587
1.198.A.	10N.286.54.F4	2	<i>Vibrio cyclitrophicus</i>	1	44,343
1.198.B.	10N.286.54.F4	2	<i>Vibrio cyclitrophicus</i>	1	44,343
1.199.A.	10N.286.55.C10	2	<i>Vibrio cyclitrophicus</i>	1	48,312
1.199.B.	10N.286.55.C10	2	<i>Vibrio cyclitrophicus</i>	1	48,312
1.200.O.	10N.286.55.E1	2	<i>Vibrio splendidus</i>	4	59,297
1.201.B.	10N.286.55.F1	2	<i>Vibrio cyclitrophicus</i>	7	50,506
1.202.O.	10N.222.45.E8	1	<i>Vibrio cyclitrophicus</i>	17	31,951
1.204.O.	10N.222.46.F12	1	<i>Vibrio cyclitrophicus</i>	11	44,168
1.205.O.	10N.222.51.A7	1	<i>Vibrio tasmaniensis</i>	7	58,011
1.206.O.	10N.222.51.B10	1	<i>Vibrio breoganii</i>	6	37,551
1.207.B.	10N.222.51.C2	1	<i>Vibrio breoganii</i>	13	55,794
1.208.B.	10N.222.52.A7	1	<i>Vibrio splendidus</i>	8	49,086
1.209.O.	10N.222.52.B2	1	<i>Vibrio breoganii</i>	8	43,503
1.210.O.	10N.222.52.C2	1	<i>Vibrio crassostreae</i>	9	48,224
1.211.A.	10N.222.52.F11	1	<i>Vibrio tasmaniensis</i>	9	37,169
1.211.B.	10N.222.52.F11	1	<i>Vibrio tasmaniensis</i>	9	37,169
1.213.O.	10N.222.54.F10	1	<i>Vibrio crassostreae</i>	1	42,443
1.214.O.	10N.222.54.F11	1	<i>Vibrio cyclitrophicus</i>	9	44,205
1.215.A.	10N.222.54.F7	1	<i>Vibrio splendidus</i>	10	80,834
1.215.B.	10N.222.54.F7	1	<i>Vibrio splendidus</i>	10	80,834
1.216.O.	10N.222.55.C12	1	<i>Vibrio crassostreae</i>	1	41,359
1.217.O.	10N.261.45.A1	1	<i>Vibrio sp. 10N-1</i>	17	31,696
1.219.O.	10N.261.45.E2	1	<i>Vibrio sp. 10N-1</i>	17	31,776
1.223.O.	10N.261.48.A9	1	<i>Vibrio tasmaniensis</i>	7	49,535

Table 3 (panel 6 of 7). Catalog of Nahant Collection phages and isolation hosts.

Phage ID	Strain ID	PFU 15mL ⁻¹	Hsp60-based population assignment	Phage Genome Cluster (Thesis)	Genome Length
1.224.A.	10N.261.48.B1	1	<i>Vibrio breoganii</i>	5	71,915
1.225.O.	10N.261.48.B7	1	<i>Vibrio splendidus</i>	7	52,615
1.226.O.	10N.261.48.E5	1	<i>Vibrio breoganii</i>	6	35,655
1.228.O.	10N.261.49.C1	1	<i>Vibrio breoganii</i>	6	36,474
1.231.O.	10N.261.49.F8	1	<i>Vibrio splendidus</i>	1	44,527
1.232.O.	10N.261.51.E11	1	<i>Vibrio tasmaniensis</i>	1	56,736
1.233.A.	10N.261.51.E6	1	<i>Vibrio breoganii</i>	6	36,823
1.233.B.	10N.261.51.E6	1	<i>Vibrio breoganii</i>	6	36,823
1.235.O.	10N.261.52.B2	1	<i>Vibrio kanaloae</i>	8	47,017
1.236.O.	10N.261.52.C4	1	<i>Vibrio cyclitrophicus</i>	1	42,646
1.237.A.	10N.261.52.C5	1	<i>Vibrio tasmaniensis</i>	4	60,222
1.237.B.	10N.261.52.C5	1	<i>Vibrio tasmaniensis</i>	4	60,160
1.238.A.	10N.261.52.F10	1	<i>Vibrio breoganii</i>	5	70,432
1.238.B.	10N.261.52.F10	1	<i>Vibrio breoganii</i>	5	70,467
1.239.O.	10N.261.52.F6	1	<i>Vibrio breoganii</i>	6	38,618
1.240.O.	10N.261.52.F8	1	<i>Vibrio breoganii</i>	6	36,710
1.242.O.	10N.261.54.B2	1	<i>Vibrio breoganii</i>	6	36,910
1.243.O.	10N.261.54.B5	1	<i>Vibrio cyclitrophicus</i>	1	48,414
1.244.A.	10N.261.54.C3	1	<i>Vibrio crassostreae</i>	3	159,885
1.245.O.	10N.261.54.C7	1	<i>Vibrio breoganii</i>	5	71,702
1.246.O.	10N.261.54.E10	1	<i>Vibrio crassostreae</i>	7	47,133
1.247.A.	10N.261.54.E12	1	<i>Vibrio splendidus</i>	9	43,896
1.247.B.	10N.261.54.E12	1	<i>Vibrio splendidus</i>	9	43,833
1.248.O.	10N.261.54.F1	1	<i>Vibrio cyclitrophicus</i>	1	48,261
1.249.A.	10N.261.55.B9	1	<i>Vibrio cyclitrophicus</i>	15	10,611
1.249.B.	10N.261.55.B9	1	<i>Vibrio cyclitrophicus</i>	15	10,611
1.250.O.	10N.261.55.E11	1	<i>Vibrio tasmaniensis</i>	4	60,036
1.251.O.	10N.261.55.E5	1	<i>Vibrio tasmaniensis</i>	4	59,649
1.253.O.	10N.286.45.B12	1	<i>Vibrio cyclitrophicus</i>	8	47,008
1.254.O.	10N.286.45.C8	1	<i>Vibrio kanaloae</i>	11	32,681
1.255.O.	10N.286.45.F1	1	<i>Vibrio kanaloae</i>	3	159,885
1.256.O.	10N.286.45.F8	1	<i>Vibrio cyclitrophicus</i>	1	48,118
1.257.O.	10N.286.46.A4	1	<i>Vibrio splendidus</i>	11	32,371
1.259.O.	10N.286.48.F4	1	<i>Vibrio splendidus</i>	16	28,145
1.261.O.	10N.286.51.A7	1	unmapped	5	71,930
1.262.O.	10N.286.51.A9	1	<i>Vibrio breoganii</i>	8	47,635
1.263.A.	10N.286.51.B1	1	<i>Vibrio cyclitrophicus</i>	1	49,577
1.263.B.	10N.286.51.B1	1	<i>Vibrio cyclitrophicus</i>	1	49,758
1.264.O.	10N.286.51.F2	1	<i>Vibrio splendidus</i>	8	47,739
1.265.O.	10N.286.52.F6	1	<i>Vibrio cyclitrophicus</i>	1	43,567
1.266.O.	10N.286.52.F9	1	<i>Vibrio breoganii</i>	6	34,788
1.267.O.	10N.286.54.A1	1	<i>Vibrio splendidus</i>	4	59,414
1.268.A.	10N.286.54.A11	1	<i>Vibrio splendidus</i>	4	59,297
1.268.B.	10N.286.54.A11	1	<i>Vibrio splendidus</i>	4	59,297
1.269.O.	10N.286.54.A6	1	<i>Vibrio splendidus</i>	4	59,738
1.270.A.	10N.286.54.A8	1	<i>Vibrio splendidus</i>	4	59,294
1.270.B.	10N.286.54.A8	1	<i>Vibrio splendidus</i>	4	59,294

Table 3 (panel 7 of 7). Catalog of Nahant Collection phages and isolation hosts.

Phage ID	Strain ID	PFU 15mL ⁻¹	Hsp60-based population assignment	Phage Genome Cluster (Thesis)	Genome Length
1.271.A.	10N.286.54.B4	1	<i>Vibrio splendidus</i>	4	59,297
1.271.B.	10N.286.54.B4	1	<i>Vibrio splendidus</i>	4	59,297
1.272.O.	10N.286.54.C4	1	<i>Vibrio splendidus</i>	4	59,297
1.273.O.	10N.286.54.C7	1	<i>Vibrio cyclitrophicus</i>	1	48,145
1.274.O.	10N.286.54.E1	1	<i>Vibrio splendidus</i>	4	59,409
1.275.O.	10N.286.54.E11	1	<i>Vibrio cyclitrophicus</i>	9	37,915
2.275.O.	10N.286.54.E11	1	<i>Vibrio cyclitrophicus</i>	3	348,971
1.276.O.	10N.286.54.E4	1	<i>Vibrio cyclitrophicus</i>	1	44,334
1.277.A.	10N.286.54.E7	1	<i>Vibrio splendidus</i>	4	59,297
1.277.B.	10N.286.54.E7	1	<i>Vibrio splendidus</i>	4	59,297
1.278.O.	10N.286.54.E8	1	<i>Vibrio cyclitrophicus</i>	1	49,231
1.280.O.	10N.286.54.F6	1	<i>Vibrio splendidus</i>	4	59,297
1.281.O.	10N.286.54.F7	1	<i>Vibrio splendidus</i>	4	59,297
1.282.A.	10N.286.54.F8	1	<i>Vibrio splendidus</i>	4	59,294
1.283.A.	10N.286.55.A1	1	<i>Vibrio splendidus</i>	4	59,530
1.283.B.	10N.286.55.A1	1	<i>Vibrio splendidus</i>	4	59,530
1.283.C.	10N.286.55.A1	1	<i>Vibrio splendidus</i>	4	59,530
1.284.A.	10N.286.55.A5	1	<i>Vibrio cyclitrophicus</i>	1	45,648
1.285.O.	10N.286.55.C12	1	<i>Vibrio kanaloae</i>	4	59,437
1.286.O.	10N.286.55.C4	1	<i>Vibrio cyclitrophicus</i>	1	49,131
1.287.O.	10N.286.55.C7	1	<i>Vibrio cyclitrophicus</i>	1	45,449
1.289.A.	10N.286.55.E8	1	<i>Vibrio cyclitrophicus</i>	1	44,529
1.291.O.	10N.286.55.F6	1	<i>Vibrio cyclitrophicus</i>	1	43,662
1.293.O.	10N.261.52.E4	4	unmapped	8	48,275

Chapter 4

Demographics of lytic viral infection of coastal ocean *Vibrio*

ABSTRACT

Viruses can usually be found in great abundance wherever there are microbial cells. The extent of bacterial death that is implied by the presence of viruses suggests that they play an important role in the ecology and evolution of microbial systems. The high diversity of both viral and bacterial genotypes in the environment, however, poses a challenge to our understanding of how these interactions are structured. Differences in host distributions and lifestyles in the environment can be expected to provide different selective regimes with respect to optimization of phage life history strategy, perhaps selecting for particular trade-offs in burst size, host range, and virion robustness, depending on the distribution of hosts in space and time and the likelihood of encountering those hosts in clonal or diverse aggregations. Thus, evaluation of phage-host interactions in the context of ecologically differentiated but closely related hosts is a powerful tool for resolving these patterns. Here, we provide a quantitative representation of a natural host-virus interaction network using marine *Vibrio* and their phages as a model system. This model system provides a platform for assignment of newly isolated environmental strains to previously identified and well described populations that have been shown to be ecologically differentiated. Using this system we provide quantitative estimates of strain-level predation susceptibility and predator load. Given the potential for host ecology and phage genome types to structure these interactions, we evaluate infection profiles in light of these elements. We find that although predator loads are highly skewed, with the majority of strains only susceptible to very few phages, at least 20% of strains are sensitive to co-occurring phages. Further, we observe differences in phage infection regimes that reflect host population structure. These differences include higher incidence of infection for some populations than others and associations of particular phage genome groups with particular host populations. We find that phage genomes are of two main types, nontailed viruses and tailed viruses, which are unconnected on the basis of shared protein clusters. Within the tailed viruses there are multiple phage groups which, though they all share at least some protein clusters in common with one of the other tailed groups, are also distinctive in their protein composition. Phage groups defined on the basis of distinctive protein cluster content show little high-identity sequence similarity with other phage groups, suggesting that they are cohesive and persistent, as members were isolated over multiple time points. Infection profiles of phages within phage genome groups are also generally similar and range from highly specific to broad host range. The abundance of highly specific phages in this

collection, and evidence for recovery of multiple highly similar hosts, suggests that “Kill the Winner”-type highly specific interactions may be common in the environment. However, the recovery of multiple isolates of a class of broadly-infecting non-tailed viruses, which are thought to be highly abundant in the environment but are underrepresented in culture collections, suggests that additional dynamics may also be important in marine systems.

INTRODUCTION

The transformations and cycling of carbon in the surface oceans determine the extent to which organic matter is available for consumption by higher trophic levels, both in the surface oceans and in the deep. Heterotrophic bacteria are abundant in the oceans and play a unique role in these transformations as they are the primary means for the return of carbon from the dissolved to the particulate pool that is accessible by higher trophic levels (Azam et al., 1983). The bacterially-mediated remineralization of detrital particulate carbon also liberates inorganic nutrients required by phytoplankton and thus plays a critical supportive role in ongoing fixation of carbon by autotrophs. The abundance and activities of heterotrophic bacteria are thus expected to have ecosystem-level relevance in the global oceans. Two main groups of organisms, protists and viruses, prey on bacteria in the surface oceans and the impacts of their predation are opposite in effect. Consumption of bacteria by protists results in incorporation of carbon into a food-web of increasingly larger predators and thus directly supports higher trophic levels and the removal of carbon from its local site of fixation. In contrast, the infection of bacteria by lytic viruses (bacteriophages, or phages) results in the lysis of cells and the return of carbon from the particulate pool back into the local dissolved pool (Fuhrman, 1999; Suttle, 2007; Wilhelm and Suttle, 1999).

The loss of bacteria to protist and viral predation has been shown to be approximately equal in marine systems but the dynamics of this loss are expected to differ due to the distinct nature of predation by these two groups. Though protists are not purely generalist in their consumption of bacteria, and can have differential impacts on different groups, they are far more generalist than viral predators, which readily discriminate bacteria at the strain level (Pernthaler, 2005). Viruses are present at concentrations of up to 10^7 mL in the coastal oceans, generally an order of magnitude higher than bacteria (Bergh et al., 1989). However, where bulk abundance of protists should convey some information about their capacity to control bacteria, such numbers

for viruses do not provide any insight into the significance of local viral predation. This difficulty arises as a consequence of the highly-specific nature of viral infections of cells, which require: first, association with an appropriate receptor; second, resistance to resident defensive systems; and third, compatibility with necessary host replication machinery. Even given the complete genomic information for both virus and host, it is currently not possible to predict whether any particular virus will be capable of infecting and lysing any particular host strain, and thus our best means for ascertaining this information currently still rest with culture-based approaches. Not only are the nature of the interactions highly specific, but both bacteria and viruses in the environment have been shown to be highly diverse (Acinas et al., 2004). In fact, highly specific viral predation has been hypothesized to be the most important driver for maintaining the exceedingly high microdiversity observed among closely related bacterial strains in the environment (Thingstad, 2000; Thingstad and Lignell, 1997; Thompson et al., 2005).

A set of models aimed at understanding processes impacting heterotrophic bacteria in the surface ocean suggested that predation by specific lytic viruses was more important for the composition of bacterial communities than for their growth rates or total abundance (Thingstad and Lignell, 1997). The mechanism for viral predation used in these models, whereby each bacterial group is controlled by an exclusive viral predator, is dubbed “Kill the Winner” and provides a means whereby competitive exclusion by any one bacterial group can be prevented by outgrowth of its specific predatory virus (Thingstad, 1976). Recent updates to this model include incorporation of a “Cost of Resistance” (KtW-CoR) that, by penalizing strains for viral resistance, provide a mechanism for explaining the observed inverse rank abundance of highly abundance resistant groups and rare susceptible groups (Våge et al., 2013a). In this KtW-CoR model the fastest growing bacteria are also those with the lowest CoR and thus are the most susceptible and consequentially the least abundant, whereas those groups that are most resistant can become quite abundant despite their lower resource competitiveness. This model recapitulates the results of the study by Waterbury & Valois that showed that: 1) most *Synechococcus* strains were resistant to infection by co-occurring phages, and 2) that the rarest strains were more susceptible to infection by phages and also better competitors for resources (Waterbury and Valois, 1993). By explicitly extending the distribution of a CoR across strains within a bacterial population or species, rather than across higher order taxa alone, this model was also used to account for the observed abundance of *Pelagibacter* phages in metagenomic

sequences (Våge et al., 2013b). Whereas the existence of phages for the highly abundant *Pelagibacter* clade was viewed by some as a challenge to KtW, the distribution of a CoR across strains within the *Pelagibacter* can account for this observation (Giovannoni et al., 2013; Våge et al., 2013b). In this case, KtW-CoR predicts that there is diversity among *Pelagibacter* strains with respect to growth rates such that rare, fast growing strains, support the bulk of *Pelagibacter* phage production. This most recent iteration of KtW-CoR also makes predictions for systems containing phages with more than one host. The reliance of KtW on exclusive pairwise interactions between phages and hosts has been a point of weakness for this theory in the face of increasing numbers of demonstrations of multihost phages in the marine environment (Holmfeldt et al., 2007; Matsuzaki et al., 1992). In this newest iteration of KtW-CoR, which incorporates nested phage infection, host ranges are hierarchically overlapping such that there are generalist phages with low adsorption coefficients infecting all hosts and progressively less generalist phages with progressively higher adsorption coefficients such that the most highly resistant host is only infected by the most specialized phage. In this model specialist phages are predicted to dominate and generalist phages to be less abundant (Vage et al., 2013). However, how well this truly reflects marine systems is unclear because, despite the evidence for marine phages with a variety of host ranges, the relative prevalence of phages with different host ranges is not known because of the lack of quantitative studies assessing host ranges of phages on co-occurring hosts (Holmfeldt et al., 2007; Våge et al., 2013a).

Current views of the nature of interactions between phages and hosts in the marine environment are largely based on studies that either use enrichments to recover phages for specific hosts or that use allochthonous hosts, or both. These approaches confound the assessment of the *in situ* dynamics of phage interactions because they can neither quantitatively represent the relative abundance of different classes of phages, nor assess the incidence of predation on resident bacterial strains. Nonetheless, studies using such approaches have provided significant insight into the nature of interactions in that they have shown that phages for marine strains can be recovered, that these predators are usually rare, and that the abundance of phages with particular specificities can cycle quite rapidly (Ahrens, 1971; Moebus and Nattkemper, 1981; Spencer, 1955; Spencer, 1960). Tests of host range between co-occurring hosts and enriched phages have shown that phages are often highly specific and can discriminate between otherwise phenotypically indistinguishable host strains (Hidaka, 1977; Moebus and Nattkemper,

1981, 1983). The capacity for viruses to control specific genotypes was shown by tracking the abundance of allochthonous hosts added into natural seawater. Using fluorescently labeled viruses as a tool to specifically enumerate the abundance of the added host strain it was possible to track the decline of the host, as well as the outgrowth of a resident predator phage specific to the added strain (Hennes et al., 1995).

In a series of extensive studies conducted over several decades, Karlheinz Moebus addressed the ecology of marine host-phage interactions using a variety of approaches. In a study of the susceptibility of 913 strains to viruses co-isolated from the same sampling effort (time and place) but using enrichment, he found that 23% of strains were susceptible to co-occurring viruses (Moebus, 1980). To address the abundance of phages *in situ* Moebus conducted a time series of direct platings of seawater in agar overlays, a quantitative method for enumeration of infective viruses in a sample, employing a panel of host strains that included isolates recently recovered from the same sampling site. In this way he found that concentrations of less than 2 plaque forming units (PFUs) mL^{-1} dominated, but that occasionally some strains could transiently recover phage abundance of up to 1,500 mL^{-1} . In an unparalleled effort, he, together with Nattkemper, conducted host range cross tests between 774 bacterial strains and 298 bacteriophages, of these, >300 bacteria and >250 phages were from his Atlantic Ocean sampling series and among these 28% infected a single host, 53% infected 1-10 hosts, and 19% infected ≥ 10 hosts. Unfortunately Moebus and Nattkemper did not have access to molecular methods for determining the phylogenetic relationships among hosts, though using physiologic studies they did find that the overwhelming majority of these bacteria were *Vibrio* (Moebus and Nattkemper, 1983). Reanalysis of this matrix 30 years later showed it to be highly modular, with discrete blocks of exclusively interacting hosts and phages (Flores et al., 2011; Flores et al., 2013). Whereas nestedness, the hierarchical overlap of phage host ranges, had been shown in previous studies and was suggestive of coevolutionary arms races as seen in previous lab studies, modularity was predicted to be a feature of fundamental shifts in host-compatibility that would limit cross-infections of phages among more distantly related hosts (Flores et al., 2011; Flores et al., 2013). Though such modules were detectable in the Moebus & Nattkemper matrix their association with phylogeny can unfortunately not be ascertained due to the lack of sequence data. Thus, overall, despite the highly dynamic nature of viral production in the surface oceans, culture based studies have suggested that members of interacting pairs of viruses and hosts are relatively

rare and that viruses are generally highly specific (Wommack and Colwell, 2000). The common use of enrichments and allochthonous indicator strains leave open the question of whether *in situ* dynamics have been accurately captured.

An additional feature of dynamics that should be considered in the study of *in situ* interactions between viruses and hosts is the non-homogeneous nature of the surface oceans. Despite the macroscopic homogeneity, seawater is microscopically highly structured and is rich in surfaces and nutrient patches such as transparent exopolymeric material, fecal pellets, and plumes of carbon around phytoplankton, sinking particles, and feeding zooplankton (Long and Azam, 2001; Stocker, 2012; Stocker et al., 2008; Verdugo et al., 2004). This patchiness creates the opportunity for bacterial adaptation to particular niches and raises the possibility that phage life-history optimization is also entrained to increase encounter with hosts distributed in temporally, spatially, and physiologically distinctive fashion. The consideration of ecological differentiation of hosts with respect to such microhabitats might thus be of substantial value in discerning structure in the nature of phage interactions with hosts, as well as in genome content of phages interacting with hosts associated with different microhabitats.

The importance of ecological differentiation in circumscribing gene flow has been shown for bacteria using the *Vibrio* as a model (Hunt et al., 2008; Shapiro et al., 2012). Using association with different particle sizes in seawater as a means to infer differences in ecology, it was shown that *Vibrio* isolated from the same sample of seawater exist as distinct populations with unique ecological distributions. Subsequent studies have further demonstrated the cohesiveness of these populations, have shown their differentiated ecological distributions to be persistent over time, and have found them to generally correspond to named species of *Vibrio* (Cordero et al., 2012; Hunt et al., 2008; Preheim et al., 2013; Szabo et al., 2012). The method used to assign associations was a size-based fractionation of water column communities using passage of seawater through sequentially smaller pore-size filters and recovery of *Vibrio* from these filters using selective media. Using this approach populations were identified that ranged from particle-associated, to free-living, to generalist in distribution. Subsequent studies of particle-associated populations showed that whereas some of these populations are associated with a variety of particle sizes, others are algal specialists with reduced genome sizes and reduced metabolic capabilities (Preheim et al., 2011a). Demonstration of restricted gene flow and

phenotypic differences between a free-living clade within an otherwise particle-associated population also provides strong evidence for the role of ecological associations in directly impacting the evolution of microbial populations in the environment (Hunt et al., 2008; Shapiro et al., 2012; Yawata et al., 2014). Thus, while closely related bacteria derive from the same samples of seawater they exist in different niches and experience segregated gene flow. This *Vibrio* model system provides the ideal foundation upon which to assess the potential for such ecological differentiation to be reflected in the structure of interactions with, and genomes of, co-occurring predatory phages. Further, the *Vibrio* are representative of a common strategy of survival of heterotrophic marine bacteria, that of the opportunitroph - occurring generally at low abundance but capable of rapid and substantial blooms in response to increased nutrients such as occur during phytoplankton blooms. As a result, patterns of host-phage interactions discernable for the *Vibrio* are likely to be broadly applicable to marine heterotrophic bacteria with this strategy.

Here, we quantitatively characterize the structure of interactions between heterotrophic marine bacteria and viruses, using the *Vibrio* model system. We provide quantitative estimates of the incidence of susceptibility and its variability in time, the abundance of phage predators, the prevalence of narrow and broad host range phages, and the relation of these properties to bacterial population structure and viral genome diversity. We use a culture-based approach, co-isolating *Vibrio* and phages at three times points and assaying for susceptibility. These censuses provide the basis for isolation and characterizations of phages, and evaluation of lytic infection patterns in light of host population structure and phage genomic diversity.

RESULTS

Prevalence of bacterial strains susceptible to viral predators

To assess virus-host interactions from the perspective of the host cell we assessed, for a large number of strains, the susceptibility of novel environmental isolates to co-occurring phages. A quantitative census of phage susceptibility showed that water-column *in situ* predator load is highly skewed, with the majority of strains having low phage predator concentrations. In our assay we exposed 1334 environmental isolates (>400 from each of three time points) to phage concentrates equivalent to 15mL of source-water (all methods presented in detail in

Chapter 3). We found that overall, 22% of isolates (295/1334) were susceptible to co-isolated phages present at concentrations greater than $\sim 0.067 \text{ mL}^{-1}$ total phage predators, our limit of detection [Table 1]. Of the strains that were positive, the majority (246/295, 83%) had total phage predator loads $< 1 \text{ mL}^{-1}$ and only 2 strains had phage predator loads $> 10 \text{ mL}^{-1}$, thus, the distribution of total predators per strain was highly skewed towards low abundance [Fig. 1]. Extrapolation of the rank abundance curve of plaque forming units (PFUs) per strain based on the best-fit power law curve ($y = 137.73x^{-1.357}$, $R^2 = 0.9495$) suggests that, given a lower limit of detection, the 1334th strain would still have had predator loads of $\sim 0.008 \text{ PFU mL}^{-1}$ or $\sim 8 \text{ PFU L}^{-1}$ [Fig. 1]. These findings suggest that phage production at any one time may be disproportionately the product of single host-phage pairwise interactions. Importantly, however, these results suggest that phage predators may be present for a large percentage of strains, though at quite low abundance.

Differences in incidence of viral predation across days

The proportion of strains identified as susceptible to phage predation at our limit of detection varied widely by sampling day [Table 1]. Comparison across the three sampling days show that an increase in the number of susceptible strains in the community (10%, 23%, and 35%, respectively, for the three days) were paralleled by an increase in the average number of PFUs present per susceptible strain (0.38 mL^{-1} , 0.58 mL^{-1} , 1.04 mL^{-1}) [Table 1]. Concentrations of silicate and combined nitrite and nitrate showed the same increasing trends across the three days, whereas ammonia and phosphate did not [Table 2]. The ratio of Si:DIN (combined nitrate, nitrite, and ammonium) showed an opposite, decreasing, trend of 7.3, 2.59, and 1.15, for the three days, respectively. The ratio of Si:DIN has been shown to be predictive of shifts in communities from diatom-copepod dominance to dinoflagellate dominance when they decline below 1:1, though the observed ratios did not decline below this number they did approach this threshold on the third day (Turner et al., 1998). Water temperatures, when integrated over the preceding week, showed that the latter two days belonged to warmer temperature regimes than the first day, where the minimum for the latter was 10.7°C , and the minima for the former were 14.9°C and 14.2° [Table 3]. Increase in temperature has been shown to be the strongest predictor of increase in total *Vibrio* abundance, suggesting that growth rates may have been elevated on the latter two days (Oberbeckmann et al., 2012; Takemura et al., 2014). The diversity in unique

Vibrio Hsp60 sequences was also higher on the latter two warmer days, when phage predation was greater, than on the first day [Table 1].

Differences in susceptibility to viral predation among host populations

Vibrio populations differed in the percentage of strains susceptible to phage predation at the limit of detection of the assay. Among the six best-represented populations of vibrios in the study, the *V. breoganii* had the highest average incidence of susceptibility (38% average across all days) [Table 4]. Previous studies have shown *V. breoganii* to be specialists for algal degradation and most often associated with the intermediate size fractions (Preheim et al., 2011b). Their distributions in this study were consistent with previous findings, and distinct from those of the other well-represented populations, with significant overrepresentation in the 5µm size-fraction and significant underrepresentation in the 63µm size-fraction [Table 5]. On average, across all days, the *V. cyclitrophicus*, *V. splendidus*, and *V. tasmaniensis* were similar in the total number of strains susceptible to phage predation (26%, 21%, and 27%, respectively) [Table 4]. Of these three, the *V. splendidus* and the *V. tasmaniensis* share a generalist distribution, with strains recovered equally from all size-fractions, whereas the *V. cyclitrophicus* are large particle specialists and are overrepresented in the largest size fraction and significantly underrepresented in the smaller size fractions [Table 5]; these distributions are also consistent with previous studies (Hunt et al., 2008; Preheim et al., 2011b). Finally, among the well-represented populations in this study the *V. kanaloae* and *V. crassostreae* had the lowest susceptibilities on average (12% and 11%, respectively). These two populations differed in their ecological associations, with *V. crassostreae* well represented in both of the larger size fractions and underrepresented in the smallest size fraction, and the *V. kanaloae* overrepresented in the two smaller size fractions and underrepresented in the largest size fraction. In the case of *V. kanaloae* this average low susceptibility is misleading in that it is driven by a single, especially low, susceptibility time point rather than a consistently lower susceptibility than the other populations [Table 4]. On the first large-scale isolation day over 120 strains of *V. kanaloae* with the same *Hsp60* sequence type were recovered (and distributed across multiple sample size fractions and replicates) and only 1 of these was susceptible to phage predators. This suggests that there was a bloom of lineages of one particular *Hsp60* sequence type and that either the single phage-susceptible member was rare and unique or that phages were present at concentrations of $\sim 5 \times 10^{-4}$

mL⁻¹ (assuming that the strains were identical and as such represent 125 replicate attempts to expose hosts to 15mL of seawater). It thus appears that although there are differences among populations in percent susceptibility these differences are not directly ascribable to association with any particular size fraction. Further, an explicit test for a correlation between susceptibility and size fraction, using phylogenetically independent contrasts, showed no significant effect (Spearman's rho and p value, by size fraction: 0.2um, 0.005, 0.9413; 1.0um, -0.107, 0.1387; 5.0um, 0.115, 0.1101; 63um, -0.177, 0.01335).

Overall prevalence of broad and narrow host range phages

To determine the overall structure of phage infection profiles, host range profiles were determined for phages isolated from the large scale quantitative censuses. One phage was randomly selected from each of the host strains on which plaques were observed and these phages were purified and assayed against a panel including all hosts susceptible to phage on any of the three days. The resultant matrix consisted of 260 unique phages derived from independent plaques and 277 host strains; of the 72,020 possible interactions 1,573 were positive [Table 7, Fig. 2]. Of 253 phages infecting *Hsp60*-sequenced hosts we found that highly specific phages were well represented, with 46% infecting only a single host and 71% infecting only a single *Hsp60* sequence type [Table 6]. A variety of host range profiles were detected, ranging from infection of single host strains to infection of many host strains in up to 7 populations [Fig. 3]; infection profiles will be discussed in greater detail in the context of phage genome groups following their introduction in the subsequent sections.

Phage genomes cluster into groups with their own unique protein sets

To assess the relation of phage genome type to infection profile all phages were genome sequenced and assigned to groups on the basis of shared protein clusters. Briefly, open reading frames (ORFs) were identified using Prodigal (Hyatt et al., 2010), ORF proteins were clustered using OrthoMCL (Fischer et al., 2011; Li et al., 2003), with a requirement for minimum 75% coverage of the longer sequence, and an e-value cut-off of 10⁻⁵, this yielded 3,331 protein clusters containing at least 2 members each. Whole genomes were then clustered on the basis of shared protein clusters using the FT ClustNSee clustering algorithm in Cytoscape (Shannon et al., 2003; Spinelli et al., 2013). This produced 18 groups, within which phages had greater

overlap in protein cluster content than they did with members in any other group [Fig. 4]. Of the 18 groups, two (G10 and G18) contained only single members considered in the host range assay (as defined by origin from different plaques during the initial isolation and lytic infection of hosts included in the host panel of the host range assay) [Table 8]. A striking feature revealed by the clustering of phage genomes is that the bulk of the collection is interconnected on the basis of shared proteins, but a single group of phages is entirely disconnected [Fig. 4]. The larger interconnected mass of genomes are all members of the tailed viruses, the Caudovirales, whereas the disconnected phage group (G15) has sequence homology with two groups of nontailed viruses, the *Tectiviridae* and the *Corticoviridae* [Table 9, Fig. 5]. Assignment to the Caudovirales is based either on electron microscopy (EM) of phage group members or by sequence homology to known Caudovirales in cases where EM was not conducted [Table 9]. The Caudovirales are comprised of the *Myoviridae*, *Siphoviridae*, and *Podoviridae*, which are morphologically defined and distinguished on the basis of their tail structure, which is long and contractile in the first case, long and noncontractile in the second, and diminutive in the third. Whereas the majority of cultured dsDNA marine phages are Caudovirales, the nontailed dsDNA viruses are underrepresented, with only 3 genera named in contrast to over 30 genera of Caudovirales.

Phage genome groups are cohesive, with little genetic exchange

To determine the extent of genetic similarity among phage groups, all genomes were compared for shared nucleotide sequences of greater than 500bp at 99%, 95%, and 75% similarity. At the 99% similarity cutoff there were no shared sequences between genomes of different phage groups, this provides some confidence in the methods used to define the clusters in that it suggests that they have coherence beyond a common subset of protein clusters (Fig. 6 panel A). At the 95% similarity cutoff only 4 groups of phages emerged as being linked: G1, G4, G6, and G7 (Fig. 6 panel B). At the level of 75% identity the majority of genome pairs sharing nucleotide sequence >500bp were all *Siphoviridae* or unclassified, though there was a single connection between a podoviral genome and a siphovirus (Fig. 6 panel C, Table 9); these results also suggests that the unclassified phages are likely to be genomically siphovirus-like.

Genome groups differ in infection profiles

Evaluation of the host range matrix in light of phage genome groups reveals that infection profiles of phages within groups are generally consistent. Several classes of host range can be identified, including groups with members that: infect multiple strains within many populations (G10, G15), infect multiple diverse strains within a population (G3, G5, G8, G13, G17), and infect few or highly similar hosts within a population (many groups) [Fig. 7, Fig.8].

Broad host range phages. Phages with the largest numbers of hosts, and infecting the largest numbers of populations, fall into only 2 genome groups representing 8% of all phages in the host range assay and 39% of all infection in the entire matrix [Table 6, Fig. 9]. Of these two groups, one is a putative siphovirus (with many proteins having significant BLAST similarity to the siphovirus *Vibrio* phage 1 [Table 9]) with a sequence length of ~80, 834bp, the other group is the well-represented nontailed virus-like G15 which have genome sizes of ~10kb and are prevalent in the collection, with 18 sequenced representatives deriving from unique plaques. These findings suggest that broad host range is itself not necessarily predictive of prevalence, as the G10 phage infected 24/277 host strains in the host range assay panel and the G15 phages infected, on average ~28 strains, and they were markedly different in overall abundance in the collection [Table 8]. It is worth considering that perhaps the G10 phage and G15 phages differ in strategy in a way that impacts their relative abundance. If G10, which was only recovered twice (with one of these phages negative in the subsequent host range assay), infects primarily strains that are also susceptible to many other phages, then it would always be present with many other plaques and therefore be less likely to be recovered during random plaque selection. This is possible given that when this phage was recovered it was selected from a host with only a single plaque, however if this scenario were correct it would be expected that the host range matrix would show high rates of co-infection by other phages on hosts infected by the G10 phage and this is not the case, thus it seems likely that this type of phage was indeed less prevalent than other phage types recovered in the study. In contrast, hosts from which G15 phages were recovered had between 1 and 49 plaques, and on average ~18, so it is likely that they were actually present in great abundance and improbable that they would have been recovered so many times if they were present as only a single or few plaques among the total. Similarly to the G10 phage, inspection of the infection profile of other phages on the hosts of G15 phages does

not suggest that they represent generally broadly susceptible hosts also infected by many other phages [Fig. 2]. Notably though, regarding potential general differences in strategy, the G10 phage derives from the early, cool, less diverse collection day, and only one of the G15 phages does. This raises the possibility that their differential recovery is in some way connected to the hypothesized differences in host growth conditions between the cooler and warmer days, a hypothesis which could be tested directly by comparing life history parameters of these two groups when infecting their shared hosts at different temperatures *in vitro*. The difference in G15 abundance across the three sampling days (1, 6, and 11, across the three days) correlates with an increase in the total number of infections of phages and hosts within each day (74, 282, and 694, for the first, second, and third co-isolation days, respectively) and strongly suggests that the presence of particular phage groups can significantly alter the properties of phage-host interaction networks.

Despite the fact the G15 phages were isolated on different days and do differ to some extent in host range, most of these phages overlap in infection on the same set of hosts. This suggests that there is specific factor whose presence (or absence) significantly increases the probability of infection by these phages. It is known that phages that share sequence homology with this group, the PR4/ PRD1-like family of tectiviruses, are dependent on the presence of a susceptibility-conferring plasmid [Table 9]. The observations of host range suggest that that is likely also to be the case for these phages, and that therefore this putative plasmid also has a wide distribution among populations of the *Vibrio*. Finally, the G15 phages have the unusual characteristic of delayed plaque formation, such that many of them do not form plaques within the first several days of direct plating, but instead can take up to at least 7-21 days to be present. This suggests either that they require some physiologic change in the host condition before they can enter or, that they form a semi-stable association with the host early on and induce out only after some extended period of time, perhaps in association with host starvation or stress response.

Population-specific broad host range phages. Several phage genome groups contain members that infect multiple diverse host *Hsp60* sequence types within a population (G3, G8, G13) [Fig. 3]. This suggests the possibility of a genome chassis suited to infection of diverse backgrounds and customizable by the presence of interchangeable population-specific gene-cassettes; this is speculation but can be assessed using an alignment of the genomes to identify the presence and

function of conserved and variable elements.

Highly specific phages. Approximately half of the phages in this collection were highly specific, infecting only a single host (46%), and 71% of phages infected only single host *Hsp60* sequence type [Table 6]. Interestingly, multiple cases of “blocks” of infection could be identified in the host matrix within groups of phages that were otherwise highly specific. Such “blocks” included multiple hosts of identical *Hsp60* sequence type of which many were infected by multiple phages with the same, or nearly the same, host range profile. These “blocks” would be consistent with the bloom of a particular lineage of host and the multiple recovery of nearly identical isolates of this host and its associated phage.

Phage genome groups interact differentially with different populations

Among the six well-represented populations in this study, it is clear that the *V. breoganii* are distinct in the groups of phages that infect them. Three groups of phages (G6, G8, and G13) infect primarily *V. breoganii* [Table 10]. These groups include both population-specific broad host range phages (G8, G13) and narrow host range phages (G6). In addition, half of the G5 phages infect multiple diverse *Hsp60* types within the *V. breoganii*, whereas the other half all have identical host ranges on *V. splendidus* of identical *Hsp60* sequence type. Some members of the G5 are also observed to infect *V. splendidus*, however, whereas the multiple *V. splendidus* hosts are of identical host *Hsp60* sequence type, the hosts of G5 phages infecting within the *V. breoganii* are diverse. Phages of the G8 also infect other populations but 70% of infections for this group are within the *V. breoganii*. Of the phage groups containing members that infect multiple populations, either individually or in aggregate, several groups are also notable for their underrepresentation in the number of infections of the *V. breoganii* (G1, G4, G7, G15) [Table 10]. Interestingly, though only two groups were identified as podoviridae in this study (G5 and G8, on the basis of sequence homology [Table 9]) both of these groups are among those interacting primarily with the *V. breoganii*. This suggests the possibility that the particularities of the ecology of *V. breoganii* are well-suited to the life-history characteristics of the podoviruses.

Infection by the G15 phages, which contributed 39% of the total observed infections, also distinguished populations [Table 8, Table 10, Fig. 9]. Infections of the six well-represented populations in this study were highly skewed, with infections of the *V. breoganii* and the *V. crassostreae* underrepresented (1% and 0% of total G15 infections), infections of *V.*

cyclitrophicus, *V. splendidus*, and *V. kanaloae* similar and intermediate in number (16%, 14%, 15%, of total G15 infections respectively), and infections of the *V. tasmaniensis* overrepresented (54% of total G15 infections) [Table 8, Table 10, Fig. 8].

Although absolute numbers suggest that *V. splendidus* were disproportionately targeted by G4 phages, with 326 of 362 G4 infections, this is misleading. The vast majority of these infections, 300 of 326, were contained within a “block” of identical *Hsp60* sequence type hosts and phages with highly similar infection profiles and thus likely represent a “bloom” of this host and the multiple recovery of phages deriving from the same lineage and responding to this bloom [Fig. 2, Table 10]. Subtracting the interactions of this overrepresented block leaves the *V. splendidus* and the *V. tasmaniensis* essentially equal in the number of infections by G4 phages [Table 10]. Overall, among the well-represented populations, phages infecting multiple hosts appeared more likely to cross infect between *V. splendidus*, *V. tasmaniensis*, and *V. kanaloae* than *V. cyclitrophicus* [Table 10, Fig. 10]. Phages infecting *V. cyclitrophicus* were also more likely to infect other strains within the *V. cyclitrophicus*, though connections with other populations were more frequent than for *V. breoganii* phages [Table 10, Fig. 10]. Several additional interactions between phage groups and host populations were noteworthy, though these populations were less well represented in the collection overall. The *Vibrio sp. 10N-1* clade was strongly tied to the G17 Mu-like myovirus group (Table 9). Of the phages infecting strains within the *Vibrio sp. 10N-1* only a single phage infected strains outside this cluster, however, the majority of susceptible hosts in this clade (7/8) derive from the same day and the majority of these (6/7) derive from the same size fraction. The two clades of *Shewanella* identified as phage-susceptible were strongly associated with G12 phages, one of which also infected a single isolate of *V. breoganii* but which were otherwise exclusive to the *Shewanella* (Fig. 2).

Highly similar phage genomes recovered on multiple days

The recovery of pairs of phage genomes from different sampling days that are highly similar suggests that there is some stability in phage-host interactions. At least 5 pairs of temporally-distanced but highly similar phage genomes were identified and, of these, 3 pairs were recovered from the G6 phages, which primarily infect the *V. breoganii*. One pair of was isolated on days 261 and 286 (1.012.O and 1.098.O, respectively). Each genome was 59,851bp long and there were only 3bp different between them [Chapter 3-Table 3]. Their host ranges

were not identical, with 1.012.O infecting 3 more strains in the host panel, and this may be due to differences in titer, which was not standardized; alternatively, the 3 base pair difference between the genomes may be associated with slight difference in host range. Homology searches for the proteins encoded by the genes impacted by the sequence changes show that one is a putative tail-associated protein (ACLAME: “tail component protein” in *Salmonella enteritidis* phage SETP3, 29% identity, $e = 2 \times 10^{-38}$; NCBI: “tail protein” in *Salmonella* phageFSL-SP062, 97% query cover, 31% identity, $e = 4 \times 10^{-95}$) and the other has no similarity to proteins of known function (ACLAME: no hits; NCBI: “hypothetical protein” in *Pseudomonas*, 47% query cover, 28% identity, $e = 3 \times 10^{-26}$). Briefly, annotations of dissimilar regions were detected by: alignment of genomes with NCBI BLAST to identify regions with polymorphisms, identification of open reading frames using GeneMark, and BLAST of proteins against the ACLAMEdb of mobile genetic elements and the NCBI nr database (Altschul et al., 1990; Besemer and Borodovsky, 2005; Leplae et al., 2004; Leplae et al., 2010).

A second pair of highly similar genomes were isolated on days 261 and 286 (1.235.O and 1.253.O, respectively) though their genomes were not identical in length, with the first 47,017bp and the second 47,008bp [Chapter 3-Table 3]. These two phages were members of the G8 podoviridae, both infecting the *V. cyclitrophicus*; host ranges for these two phages were not identical, with 1.253.O infecting several more hosts than 1.235.O, but again this may be attributable to differences in titer rather than a true difference in host range. Inspection of the three regions of difference between the genomes shows one gene with an insertion of 9bp, and 2 sites with mismatches. In the first case the difference is in a protein that occurs in a variety of phages and is annotated as hypothetical (NCBI: “hypothetical protein ORF041” in *Pseudomonas aeruginosa* phage PA11, 97% query cover, 31% identity, $e = 1 \times 10^{-15}$). The second difference is in a gene annotated as tail-associated (NCBI: “putative tail appendage” in *Salinivibrio* phage CW02, 89% query cover, 33% identity, $e = 8 \times 10^{-106}$). The third is again a hypothetical but present in many phages (NCBI: “hypothetical protein ORF015, *Pseudomonas* phage PA11, 100% query cover, 49% identity, $e = 2 \times 10^{-72}$).

Group 6 phages yielded 3 pairs of nearly identical phage genomes from separate days. The first pair, 1.064.O and 1.108.O were both 36,584bp, differed in only a single base, and both infected the same two hosts (10N.222.52.A4 and 10N.261.52.E2), though these hosts have not

been identified to specific populations [Chapter 3-Table 3]. The gene affected by the polymorphism has no significant similarity to other sequences in NCBI nr database. The second pair, 1.116.O and 1.172.O are both 36,314bp and both sequences are identical; both infect the same 2 *V. breoganii* host strains [Chapter 3-Table 3]. The third pair, 1.147.O and 1.162.O, are both 37,360bp are both also identical, and both infect the same 2 hosts, one of which is *V. breoganii* (10N.261.48.E3) and the other of which has not been mapped to a population (10N.286.49.E9) [Chapter 3-Table 3].

DISCUSSION

In most regards, the results of this study with respect to bulk properties of incidence of susceptibility, abundance of phage predators, and prevalence of phages with narrow host ranges, were surprisingly similar to those of previous studies using enrichment and allochthonous host approaches. Direct plating of viral concentrates was used to assess the prevalence of susceptibility to co-occurring phages in *Vibrio*, though this differed to some extent across sampling days the grand average was 22% susceptible. In a previous large-scale study of susceptibility to co-occurring viruses using enrichment approaches, 23% of strains were shown to be susceptible to co-occurring phages (Moebus, 1980). The overwhelming majority of these susceptible strains were later shown to be *Vibrio* though they were not specifically targeted or collected using selective media. In the Moebus study direct nutrient enrichment of natural communities was used to stimulate production of viruses and increase their abundances and then these enrichments were filtered to remove cells and used in direct platings on co-isolated hosts. The fact that these two studies show similar results suggests that when *Vibrio* increase in abundance they do so as a diverse group and thus enrich for diverse associated viruses. The observation in this study that on the two days with higher temperature, where *Vibrio* growth rates are expected to be higher, the diversity was greater than on the cooler day is consistent with these observations.

The recent digitization and re-analysis of the host range matrix generated by Moebus & Nattkemper for the collection of Atlantic series vibrios and their phages facilitates the natural comparison of that study with the one described here (Flores et al., 2013; Moebus, 1980; Moebus

and Nattkemper, 1981). Overall, the properties of these two matrices were strikingly similar¹ and compare as follows, with the Moebus & Nattkemper and Nahant Collection studies indicated as (MN) and (NC), respectively. The total numbers of hosts were 286 (MN) and 277 (NC), the number of phages 215 (MN) and 260 (NC), the number of interactions 1,332 (MN) and 1,573 (NC), and the connectance was 0.02 in both. This measure of connectance is far lower than the average connectance reported for phage-host matrices, which in the meta-analysis of Flores was 0.39. Though in their meta-analysis they indicate that connectance (the percent of positive interactions out of the total possible) is not correlated with the number of species it is clear that in different size matrices with the same infection characteristics the size of the matrix matters, such that smaller matrices have larger connectance. For example, in a 2x2 matrix, where each phage infects only its own host, the connectance is 0.5; in a 10x 10 matrix where each phage infects only its own host, the connectance is 0.1. For this reason, the direct comparison of connectance across matrices does not seem useful unless in some way corrected for the size. If the MN and NC matrices had the property of each phage infecting only its own host and no others, their connectances would be expected to be 0.004 and 0.003, respectively, approximately 5-7-fold lower than observed. This suggests that in both cases they were more connected than would be predicted for pure mutually-exclusive pairwise interactions between phage host pairs. This is evidenced by the perhaps more useful measure of interactions per host and per phage; for these two matrices these values were 4.66 (MN) and 5.68 (NC) for the host mean interactions, and 6.2 (MN) and 6.1 (NC) for the phage mean interactions.

The re-analysis by Flores identified a highly modular structure within the MN matrix, such that there were blocks of hosts and phages that were mutually exclusive and to some degree nested. Due to the lack of phylogenetic information for the MN matrix it was not possible to resolve the host relationships in these blocks, nor those of the phages. Inspection of the digitized matrix provided with the study reveals that many of the modules of the MN matrix consist of highly overlapping susceptibility and host range profiles of the hosts and phages respectively. These are highly reminiscent of the “blocks” identified in this study and that were associated with hosts of identical *Hsp60* sequence type and the phages isolated from them. There are also

¹ In the Discussion of his post hoc study to determine the taxonomy of all of the phage-host systems he had assayed in his very large host range assay, Moebus writes that: “stupendous uniformity was found in the 250 bacteria placed in group I”. The use of word “stupendous” is fairly rare in the scientific literature and this may be an artifact of translation, however, it suggests some surprise on his part. In the case of the comparison between these two matrices it is also tempting to use such words with respect to the extent of their similarity.

larger and more sparsely populated modules of interacting phages and hosts in the MN matrix that are akin to those blocks observed for the *V. breoganii* wherein there were population-level specialist phages as well narrow host range phages.

Phage infection profiles in the host range matrix conducted in this study were seen clearly to reflect population structure, with some populations more deeply segregated than others. The *V. breoganii* stand out in that the phages groups that infect them mostly contain members that only infect this population. This is in contrast to other phage groups whose members, though individually specific, infect across multiple populations. They also stand out for their higher susceptibility to phage predation and the greater consistency in percent strains susceptible at different points in time. The *V. breoganii* are known to have a distinctive ecology, with reduced motility and genome sizes; it is interesting to consider that phages interacting with this population have fewer options for integration or long-term associations (as predicted on the basis of observed genome streamlining in these hosts) and that therefore those phages infecting them are more dependent on lytic modes of replication. Finally, the *V. breoganii* stand out for their overrepresentation with respect to infection by 2 groups of putative *Podoviridae*, infections by *Podoviridae* are underrepresented on all other populations though abundant for the *V. breoganii*. Previous studies have suggested that the *Podoviridae* are highly specific and though this was largely the case for those observed here, a subset of strains within one group (G8) of putative podoviruses showed a broad host range very similar to, though not overlapping with, a group of *Siphoviridae* (G13). Closer inspection of this set is warranted to resolve whether this represents a series of mixed samples or whether in fact this group is split with respect to host range profiles. In marine bacterial host systems it is generally possible to recover all three morphotypes of the *Caudovirales* for any given host strain, though some morphotypes may be more prevalent than others. In the context of an effort to understand the conditions that select for differential success of phage morphotypes it is interesting to note that the podoviruses stand out in a mycobacterial host system. Despite being the subject of intensive efforts to recover diverse phages, with hundreds of phages isolated and characterized, apparently no podoviruses have been recovered for the mycobacteria (Hatfull et al., 2008).

Though these culture-based approaches provide the best means available for assessing these relationships these findings are conservative in that they are limited by the requirement for formation of plaques visible by eye under culture conditions not directly representative of those

in the environment. Care was taken in these studies to minimize operational limitation of recovery, by for example: using low percentage top agar to prevent loss of large viruses, adding chitin to media to stimulate expression of substrate-responsive receptors, using glycerol in the top agar to increase detection of faint plaques, and allowing for extended incubation periods to identify late-forming plaques. Nonetheless, the results as presented are limited in scope to those phage-host systems amenable to formation of plaques under the conditions employed, and the genomes presented are those that were amenable to standard procedures for recovery of dsDNA phage genomes.

CONCLUSION

This study has shown that predation by phages differs among co-occurring, closely-related, but ecologically differentiated populations. These differences were evident in the percent of strains susceptible in each population and the genomic types of phages infecting them. The evaluation of phage genomes in the context of this study has also shown that there is cohesiveness among phages within genome types and that entire phage genomes can persist and be recovered from natural environments over at least several weeks. These findings suggest that despite the great diversity of hosts and viruses there is stability, and that it is structured by interactions with specific host groups.

Despite the expectation that previous enrichment studies were perhaps biased towards recovery of narrow host range viruses this study also shows that highly specific phages are prevalent in natural systems. This supports the model of “Kill the Winner” type interactions that are generally exclusive, and several lines of evidence from this study underscore this. Firstly, though only ~22% of hosts were found to be susceptible to phages in this study, this was based on a limit of detection of 0.07 phages mL⁻¹ and it is likely that additional susceptible strains would be recovered given exposure to a greater volume of viral concentrate; this suggests that phage predators are present for many host strains and would be capable of responding given a bloom of their host. Secondly, the preponderance of identical Hsp60 sequence type *V. kanaloae* on sampling day 222, as well as the “block” infections of Hsp60 identical hosts by highly similar phages, suggests that clonal expansions of host lineages may in fact be fairly common. Given this pool of resident high diversity highly-specific phages and the observed apparent lineage expansions, it appears that “Kill the Winner” type dynamics may well be supported in natural

systems. Further studies of the temporal dynamics of such apparent blooms, as that observed for the *V. kanaloae*, can be conducted using clones isolated on that day and screening them for the expected appearance of increasing numbers of PFUs in the viral concentrates of the days following the bloom, materials for such study are available as part of the Nahant Collection (Chapter 3).

Finally, a class of non-tailed phages with hosts in multiple populations was well-represented in this study though they are generally rare in other culture collections. These phages contributed substantially to the total numbers of infections and may thus be particularly interesting with respect to their capacity to mediate gene transfer among more distantly related groups through access to “hub” strains containing the susceptibility conferring factor. The observed delayed infection phenotype suggests a capacity either for infection of stationary hosts or for establishing longer-term associations with hosts before initiating the lytic cycle. Given the evidence for the dominance of nontailed phages in surface oceans the Nahant phages representing this class provide an ideal model system for further development of our understanding of the ecology and interactions of this group of phages.

METHODS

All methods used in conducting the work presented in this Chapter are detailed in Chapter 3.

REFERENCES

- Acinas, S.G., Klepac-Ceraj, V., Hunt, D.E., Pharino, C., Ceraj, I., Distel, D.L., Polz, M.F., 2004. Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 430, 551-554.
- Ahrens, R., 1971. Untersuchungen zur Verbreitung von Phagen der Gattung *Agrobacterium* in der Ostsee. *Kieler Meeresforsch* 27, 102-112.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215, 403-410.
- Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., Thingstad, F., 1983. The ecological role of water-column microbes in the sea. *Marine ecology progress series*. Oldendorf 10, 257-263.
- Bergh, Ø., Børsheim, K.Y., Bratbak, G., Heldal, M., 1989. High abundance of viruses found in aquatic environments. *Nature* 340, 467-468.
- Besemer, J., Borodovsky, M., 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic acids research* 33, W451-W454.
- Cordero, O.X., Wildschutte, H., Kirkup, B., Proehl, S., Ngo, L., Hussain, F., Le Roux, F., Mincer, T., Polz, M.F., 2012. Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science* 337, 1228-1231.
- Fischer, S., Brunk, B.P., Chen, F., Gao, X., Harb, O.S., Iodice, J.B., Shanmugam, D., Roos, D.S., Stoeckert, C.J., 2011. Using OrthoMCL to Assign Proteins to OrthoMCL-DB Groups or to Cluster Proteomes Into New Ortholog Groups. *Current Protocols in Bioinformatics*, 6.12. 11-16.12. 19.
- Flores, C.O., Meyer, J.R., Valverde, S., Farr, L., Weitz, J.S., 2011. Statistical structure of host-phage interactions. *Proceedings of the National Academy of Sciences* 108, E288-E297.
- Flores, C.O., Valverde, S., Weitz, J.S., 2013. Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. *The ISME Journal* 7, 520-532.
- Fuhrman, J.A., 1999. Marine viruses and their biogeochemical and ecological effects. *Nature* 399, 541-548.
- Giovannoni, S., Temperton, B., Zhao, Y., 2013. Giovannoni et al. reply. *Nature* 499, E4-E5.
- Hatfull, G.F., Cresawn, S.G., Hendrix, R.W., 2008. Comparative genomics of the mycobacteriophages: insights into bacteriophage evolution. *Research in Microbiology* 159, 332-339.
- Heidelberg, J.F., Paulsen, I.T., Nelson, K.E., Gaidos, E.J., Nelson, W.C., Read, T.D., Eisen, J.A., Seshadri, R., Ward, N., Methe, B., 2002. Genome sequence of the dissimilatory metal ion-reducing bacterium *Shewanella oneidensis*. *Nature Biotechnology* 20, 1118-1123.
- Hennes, K.P., Suttle, C.A., Chan, A.M., 1995. Fluorescently Labeled Virus Probes Show that Natural Virus Populations Can Control the Structure of Marine Microbial Communities. *Applied and environmental microbiology* 61, 3623-3627.
- Hidaka, T., 1977. Detection and isolation of marine bacteriophage systems in the southwestern part of the Pacific Ocean. *Mem. Fac. Fish. Kagoshima Univ* 26, 55-62.
- Holmfeldt, K., Middelboe, M., Nybroe, O., Riemann, L., 2007. Large variabilities in host strain susceptibility and phage host range govern interactions between lytic marine phages and their *Flavobacterium* hosts. *Appl Environ Microbiol* 73, 6730-6739.
- Hunt, D.E., David, L.D., Gevers, D., Preheim, S.P., Alm, E.J., Polz, M.F., 2008. Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* 320, 1081-1085.

- Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 119.
- Lavigne, R., Darius, P., Summer, E.J., Seto, D., Mahadevan, P., Nilsson, A.S., Ackermann, H.W., Kropinski, A.M., 2009. Classification of Myoviridae bacteriophages using protein sequence similarity. *BMC microbiology* 9, 224.
- Leplae, R., Hebrant, A., Wodak, S.J., Toussaint, A., 2004. ACLAME: a CLAssification of Mobile genetic Elements. *Nucleic acids research* 32, D45-D49.
- Leplae, R., Lima-Mendez, G., Toussaint, A., 2010. ACLAME: a CLAssification of Mobile genetic Elements, update 2010. *Nucleic acids research* 38, D57-D61.
- Li, L., Stoeckert, C.J., Roos, D.S., 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome research* 13, 2178-2189.
- Long, R.A., Azam, F., 2001. Microscale patchiness of bacterioplankton assemblage richness in seawater. *Aquatic Microbial Ecology* 26, 103.
- Matsuzaki, S., Tanaka, S., Koga, T., Kawata, T., 1992. A broad-host-range vibriophage, KVP40, isolated from sea water. *Microbiology and immunology* 36, 93-97.
- Moebus, K., 1980. A method for the detection of bacteriophages from ocean water. *Helgoländer Meeresuntersuchungen* 34, 1-14.
- Moebus, K., Nattkemper, H., 1981. Bacteriophage sensitivity patterns among bacteria isolated from marine waters. *Helgoländer Meeresuntersuchungen* 34, 375-385.
- Moebus, K., Nattkemper, H., 1983. Taxonomic investigations of bacteriophage sensitive bacteria isolated from marine waters. *Helgoländer Meeresuntersuchungen* 36, 357-373.
- Oberbeckmann, S., Fuchs, B.M., Meiners, M., Wichels, A., Wiltshire, K.H., Gerds, G., 2012. Seasonal dynamics and modeling of a *Vibrio* community in coastal waters of the North Sea. *Microbial Ecology* 63, 543-551.
- Pernthaler, J., 2005. Predation on prokaryotes in the water column and its ecological implications. *Nature Reviews Microbiology* 3, 537-546.
- Preheim, S.P., Boucher, Y., Wildschutte, H., David, L.A., Veneziano, D., Alm, E.J., Polz, M.F., 2011a. Metapopulation structure of Vibrionaceae among coastal marine invertebrates. *Environ Microbiol* 13, 265-275.
- Preheim, S.P., Perrotta, A.R., Martin-Platero, A.M., Gupta, A., Alm, E.J., 2013. Distribution-based clustering: using ecology to refine the operational taxonomic unit. *Appl Environ Microbiol* 79, 6593-6603.
- Preheim, S.P., Timberlake, S., Polz, M.F., 2011b. Merging taxonomy with ecological population prediction in a case study of Vibrionaceae. *Appl Environ Microbiol* 77, 7195-7206.
- Seguritan, V., Feng, I.-W., Rohwer, F., Swift, M., Segall, A.M., 2003. Genome sequences of two closely related *Vibrio parahaemolyticus* phages, VP16T and VP16C. *Journal of Bacteriology* 185, 6434-6447.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13, 2498-2504.
- Shapiro, B.J., Friedman, J., Cordero, O.X., Preheim, S.P., Timberlake, S.C., Szabó, G., Polz, M.F., Alm, E.J., 2012. Population genomics of early events in the ecological differentiation of bacteria. *Science* 336, 48-51.
- Spencer, R., 1955. A marine bacteriophage. *Nature* 175, 690-691.
- Spencer, R., 1960. Indigenous marine bacteriophages. *Journal of Bacteriology* 79, 614.

Spinelli, L., Gambette, P., Chapple, C.E., Robisson, B., Baudot, A., Garreta, H., Tichit, L., Guénoche, A., Brun, C., 2013. Clust&See: A Cytoscape plugin for the identification, visualization and manipulation of network clusters. *BioSystems* 113, 91-95.

Stocker, R., 2012. Marine Microbes See a Sea of Gradients. *Science* 338, 628-633.

Stocker, R., Seymour, J.R., Samandani, A., Hunt, D.E., Polz, M.F., 2008. Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches. *Proc Natl Acad Sci U S A* 105, 4209-4214.

Suttle, C.A., 2007. Marine viruses--major players in the global ecosystem. *Nature Reviews Microbiology* 5, 801-812.

Szabo, G., Preheim, S.P., Kauffman, K.M., David, L.A., Shapiro, J., Alm, E.J., Polz, M.F., 2012. Reproducibility of Vibrionaceae population structure in coastal bacterioplankton. *ISME J.*

Takemura, A.F., Chien, D.M., Polz, M.F., 2014. Associations and dynamics of Vibrionaceae in the environment, from the genus to the population level. *Front Microbiol* 5, 38.

Thingstad, T.F., 1976. Connection between Bacterial-Growth and Division. *B Math Biol* 38, 415-423.

Thingstad, T.F., 2000. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnology and Oceanography* 45, 1320-1328.

Thingstad, T.F., Lignell, R., 1997. Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquatic Microbial Ecology* 13, 19-27.

Thompson, J.R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D.E., Benoit, J., Sarma-Rupavtarm, R., Distel, D.L., Polz, M.F., 2005. Genotypic diversity within a natural coastal bacterioplankton population. *Science* 307, 1311-1313.

Turner, R.E., Qureshi, N., Rabalais, N.N., Dortch, Q., Justic, D., Shaw, R.F., Cope, J., 1998. Fluctuating silicate: nitrate ratios and coastal plankton food webs. *Proceedings of the National Academy of Sciences* 95, 13048-13051.

Vage, S., Storesund, J.E., Thingstad, T.F., 2013. Adding a cost of resistance description extends the ability of virus-host model to explain observed patterns in structure and function of pelagic microbial communities. *Environ Microbiol* 15, 1842-1852.

Våge, S., Storesund, J.E., Thingstad, T.F., 2013a. Adding a cost of resistance description extends the ability of virus--host model to explain observed patterns in structure and function of pelagic microbial communities. *Environmental microbiology* 15, 1842-1852.

Våge, S., Storesund, J.E., Thingstad, T.F., 2013b. SAR11 viruses and defensive host strains. *Nature* 499, E3-E4.

Verdugo, P., Alldredge, A.L., Azam, F., Kirchman, D.L., Passow, U., Santschi, P.H., 2004. The oceanic gel phase: a bridge in the DOM-POM continuum. *Marine Chemistry* 92, 67-85.

Waterbury, J.B., Valois, F.W., 1993. Resistance to co-occurring phages enables marine synechococcus communities to coexist with cyanophages abundant in seawater. *Appl Environ Microbiol* 59, 3393-3399.

Wilhelm, S.W., Suttle, C.A., 1999. Viruses and nutrient cycles in the sea viruses play critical roles in the structure and function of aquatic food webs. *Bioscience* 49, 781-788.

Wommack, K.E., Colwell, R.R., 2000. Virioplankton: viruses in aquatic ecosystems. *Microbiology and molecular biology reviews* : MMBR 64, 69-114.

Yawata, Y., Cordero, O.X., Menolascina, F., Hehemann, J.-H., Polz, M.F., Stocker, R., 2014. Competition-dispersal tradeoff ecologically differentiates recently speciated marine bacterioplankton populations. *Proceedings of the National Academy of Sciences*, 201318943.

FIGURES AND TABLES

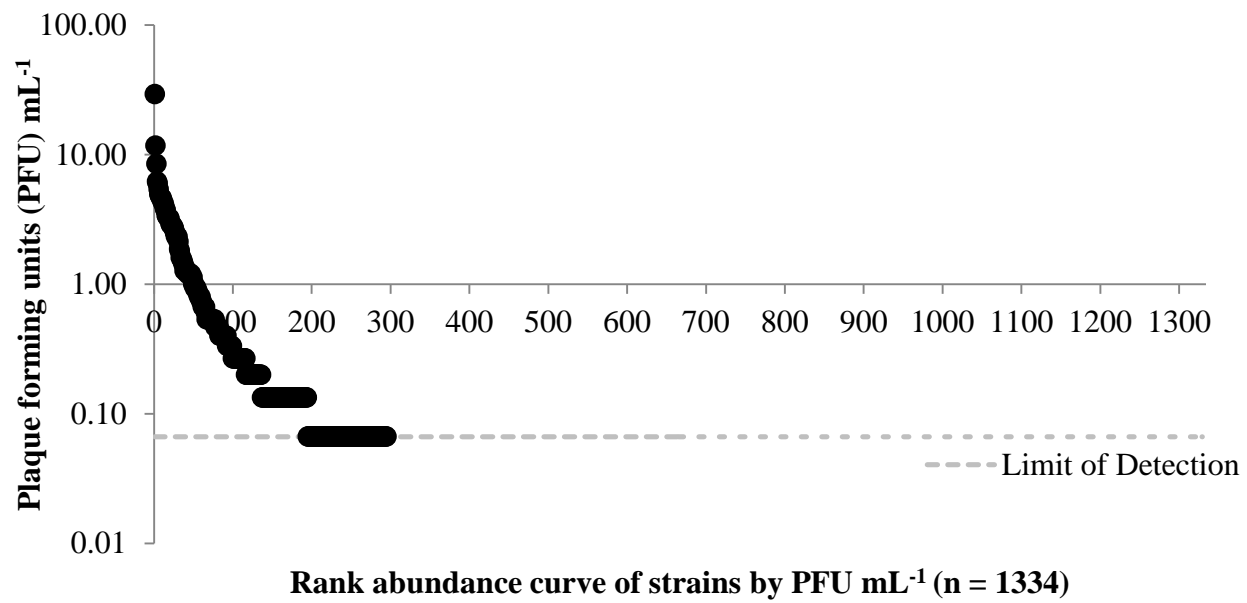


Figure 1. Phage predators per *Vibrio* strain as assayed in the large scale direct plating assay, rank abundance curve.

Figure 2. Host range matrix for assay of phages on *Hsp60*-sequenced hosts. Rows represent strains, columns represent phages, black marks indicate infection of host. Row color indicates population assignment for the majority of strains in the clade (Population legend below); column A indicates population assignment for individual strains (Population legend below); column B indicates day of strain isolation (Day legend below); column C indicates size fraction from which strain was isolated (Size Fraction legend below). Phages arrayed by genome group (genome group number indicated above the figure); phages with unknown genome group arranged according to similarity of infection profile and indicated with black block at top of matrix column; preliminary assignments of phage genome groups to morphotype indicated by color bands at top of matrix (*Myoviridae*, red; *Siphoviridae*, green; *Podoviridae*, yellow; unclassified, grey) with unknowns arranged according to infection profile similarity with group number indicated above the matrix.

Legends

Population

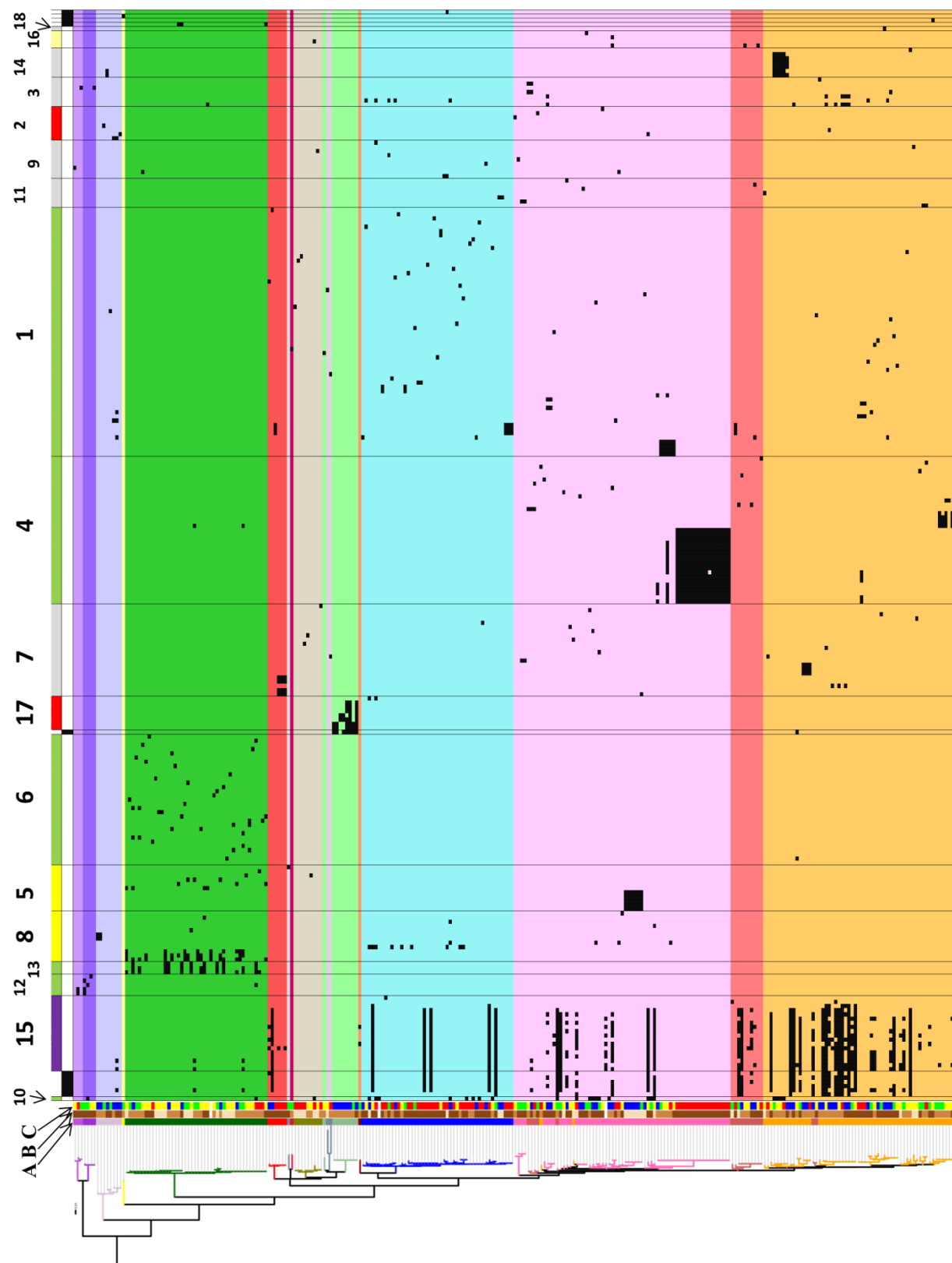
<i>Shewanella</i> sp. 9ZD10-like G2
<i>Shewanella</i> sp. 9ZD10-like G1
<i>Enterovibrio norvegicus</i>
<i>Vibrio</i> sp. F6
<i>Vibrio breoganii</i>
<i>Vibrio</i> sp. F12
<i>Vibrio splendidus</i> F11
<i>Vibrio crassostreae</i>
<i>Vibrio</i> sp. 10N-1
<i>Vibrio cyclitrophicus</i>
<i>Vibrio splendidus</i>
<i>Vibrio kanaloae</i>
<i>Vibrio tasmaniensis</i>

Isolation Day

Ordinal day 222
Ordinal day 261
Ordinal day 286

Size Fraction

0.2um
1.0 um
5.0 um
63.0 um



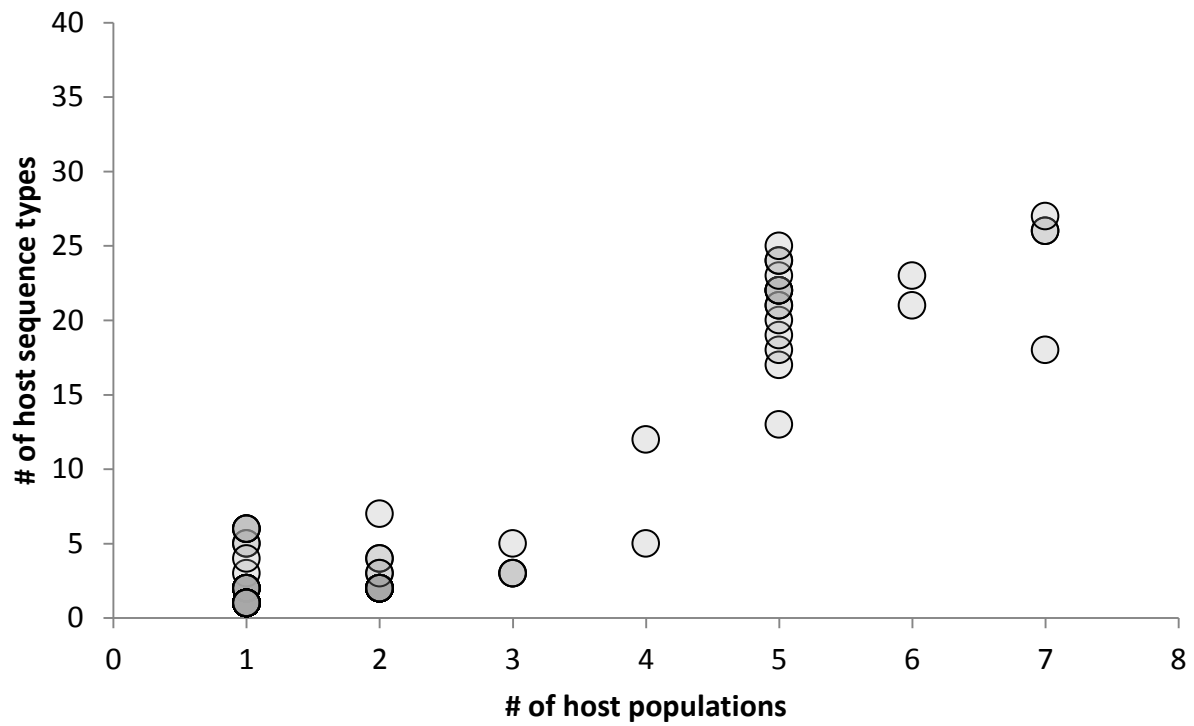


Figure 3. Phage infection profiles on hosts from all days, with respect to number of host populations and number of hosts infected. Color intensity increases with number of overlapping points ($n = 253$, excludes interactions with hosts not mapped to populations).

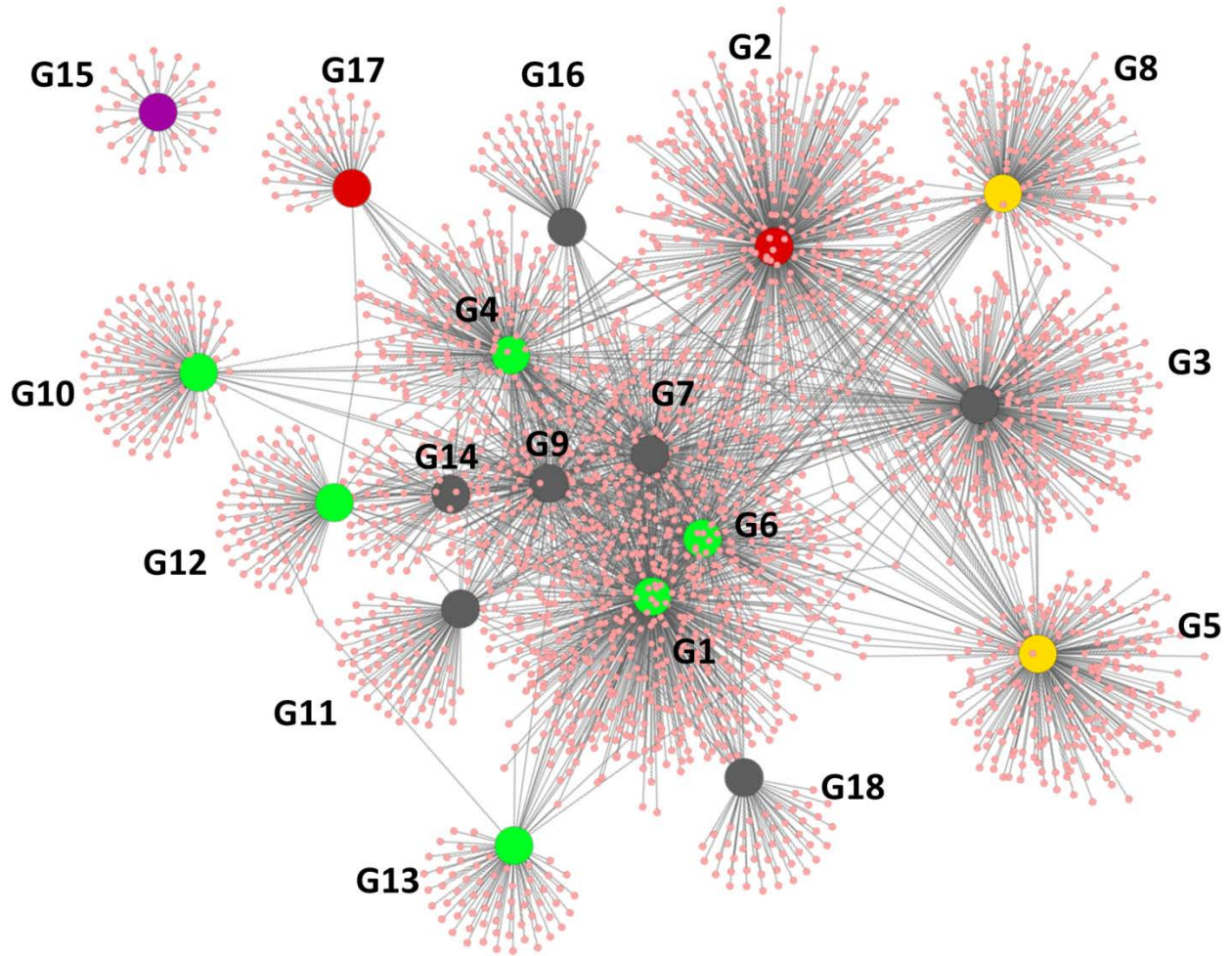
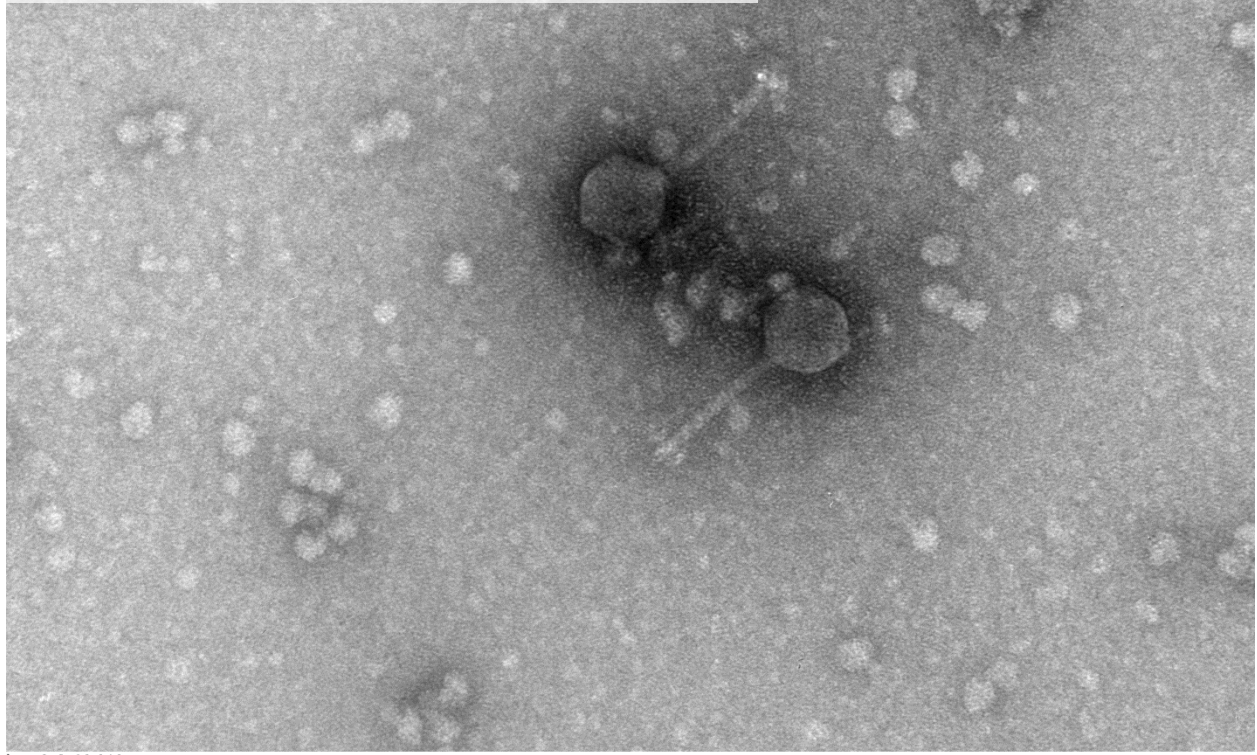


Figure 4. Phage genome clusters. Phage genome clusters (larger circles) shown linked to the protein clusters (smaller pink circles) which they share. Phage genome cluster indicates preliminary assignment to morphology (purple, *nontailed*-like; red, *Myoviridae*; yellow, *Podoviridae*; green, *Siphoviridae*; grey, unclassified).

A Siphovirus *Vibrio* phage Jenny 12G5 C1
(Phage Group 1)

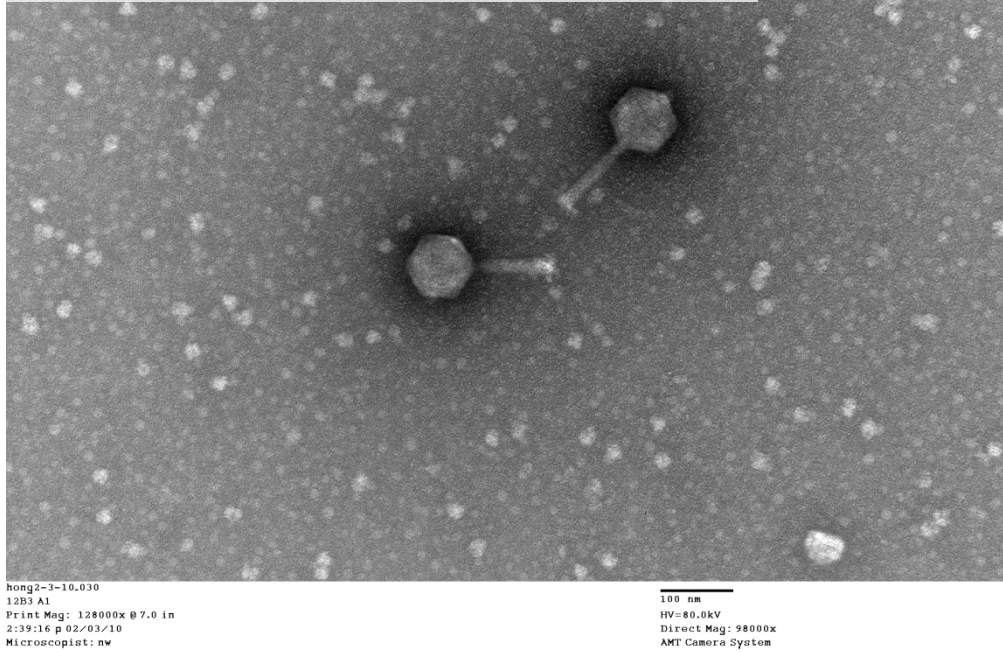


hong2-3-10.046
1265 C1
Print Mag: 197000x @ 7.0 in
3:11:11 p 02/03/10
Microscopist: nw

100 nm
HV=80.0kV
Direct Mag: 150000x
AMT Camera System

Figure 5 (Panel 1 of 9). Electron microscopy of phages. Electron microscopy by Nicki Watson.

B Myovirus *Vibrio* phage Helene 12B3
(Phage Group 2)



C Myovirus *Vibrio* phage Helene 12B3 with contracted tail.
(Phage Group 2)

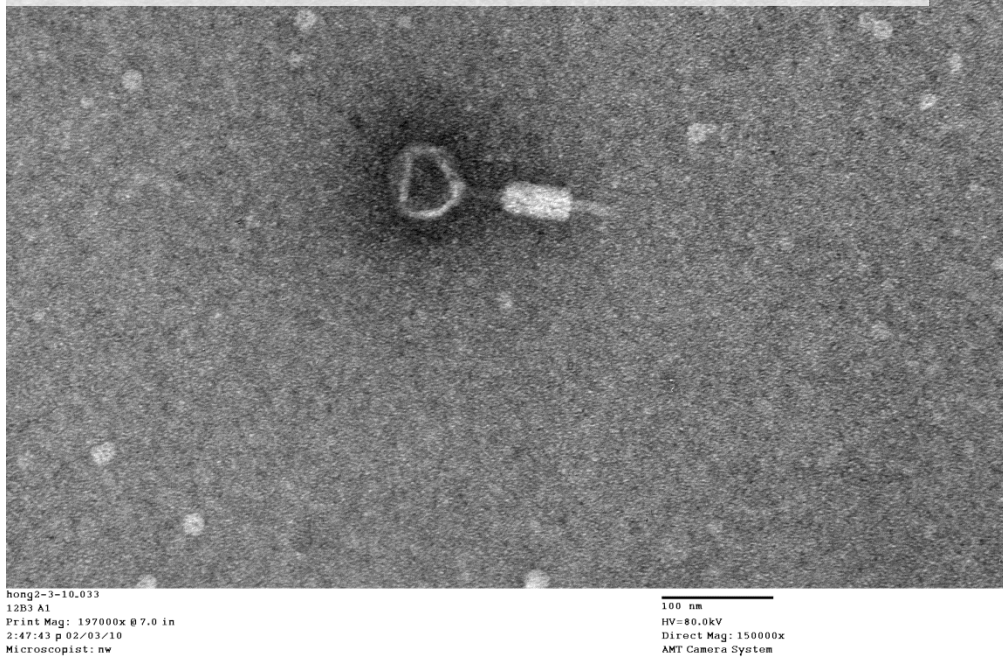
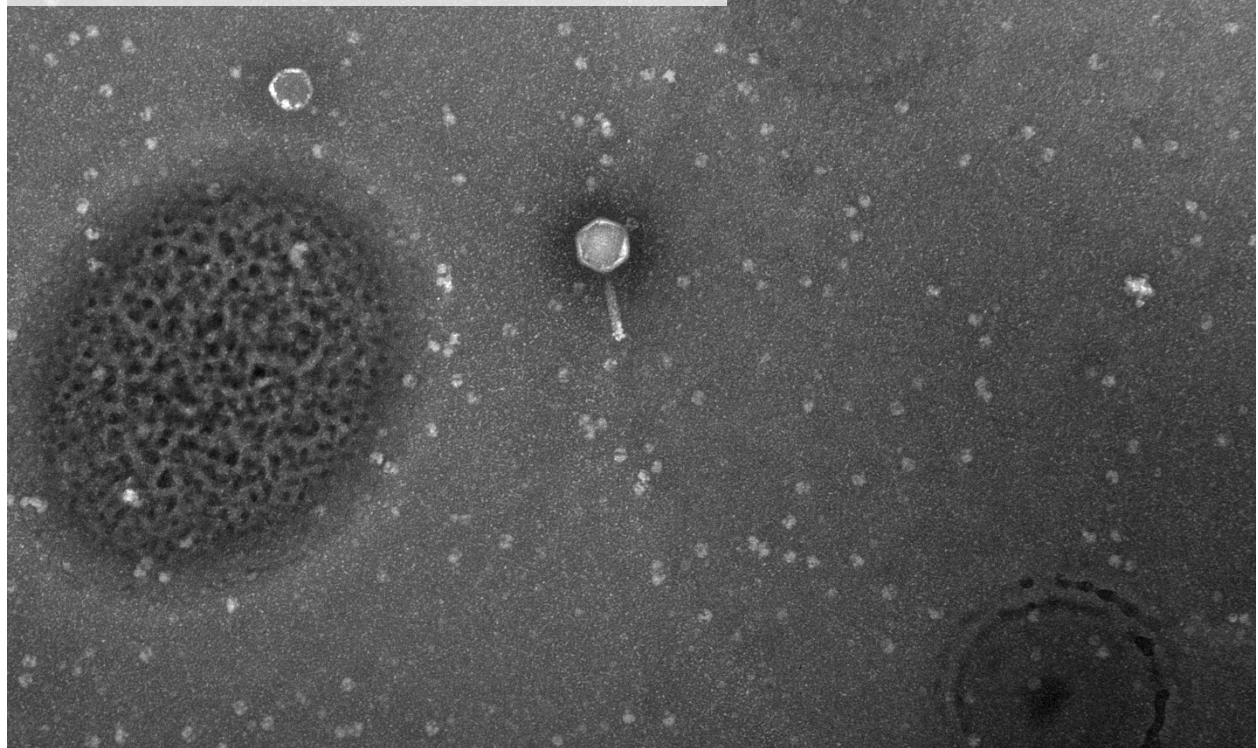


Figure 5 (Panel 2 of 9). Electron microscopy of phages. Electron microscopy by Nicki Watson.

D Myovirus Vibrio phage Eugene 12A10 C2
(Phage Group 2)

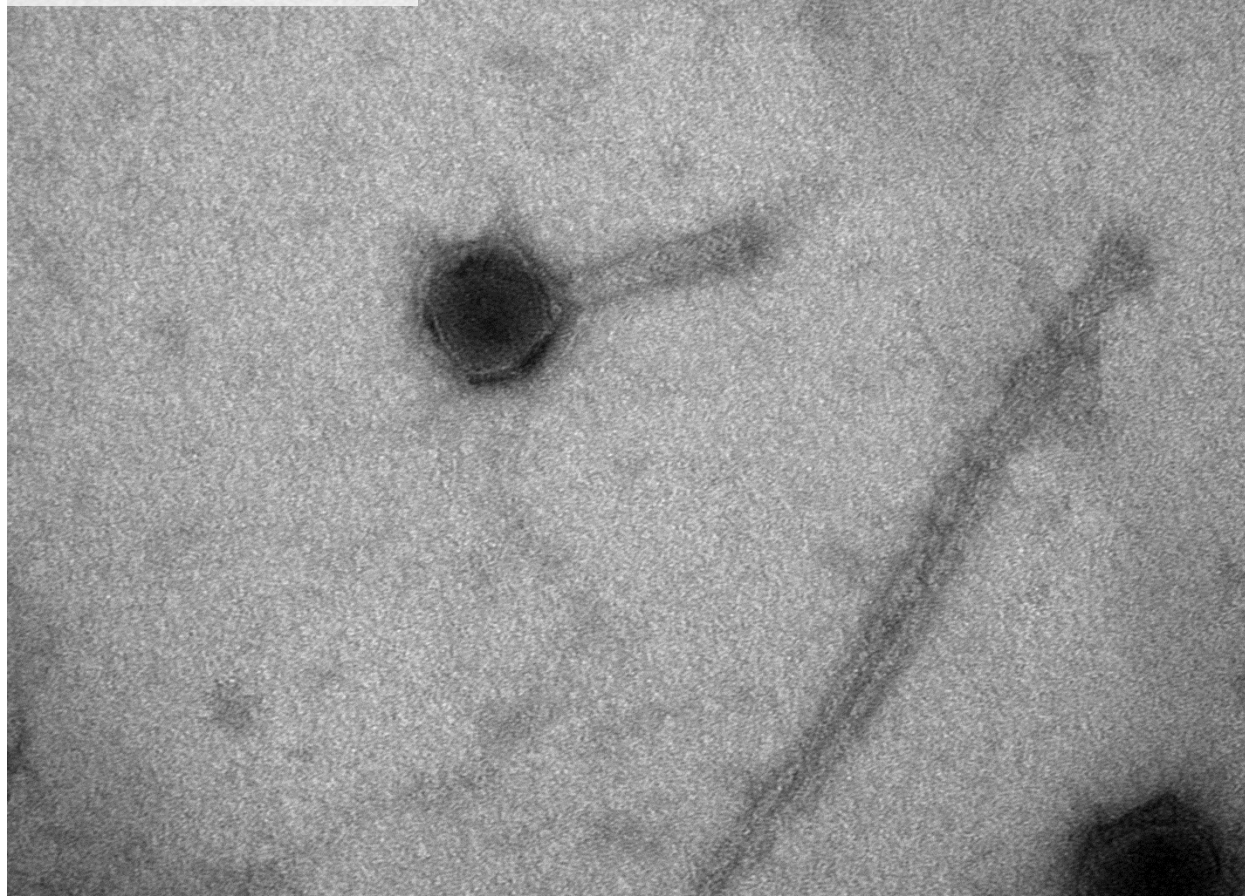


hong2-3-10.052
12A10 C2
Print Mag: 89100x @ 7.0 in
3:27:52 p 02/03/10
Microscopist: nw

100 nm
HV=80.0kV
Direct Mag: 68000x
AMT Camera System

Figure 5 (Panel 3 of 9). Electron microscopy of phages. Electron microscopy by Nicki Watson.

E **Myovirus 1.084.O**
(Phage Group 2)



katryn2-4-13_031
1.084
Cal: 0.247153 nm/pix
TEM Mode: Imaging
Microscopist: nw

100 nm
HV=80.0kV
Direct Mag: 150000x
AMT Camera System

Figure 5 (Panel 4 of 9). Electron microscopy of phages. Electron microscopy by Nicki Watson.

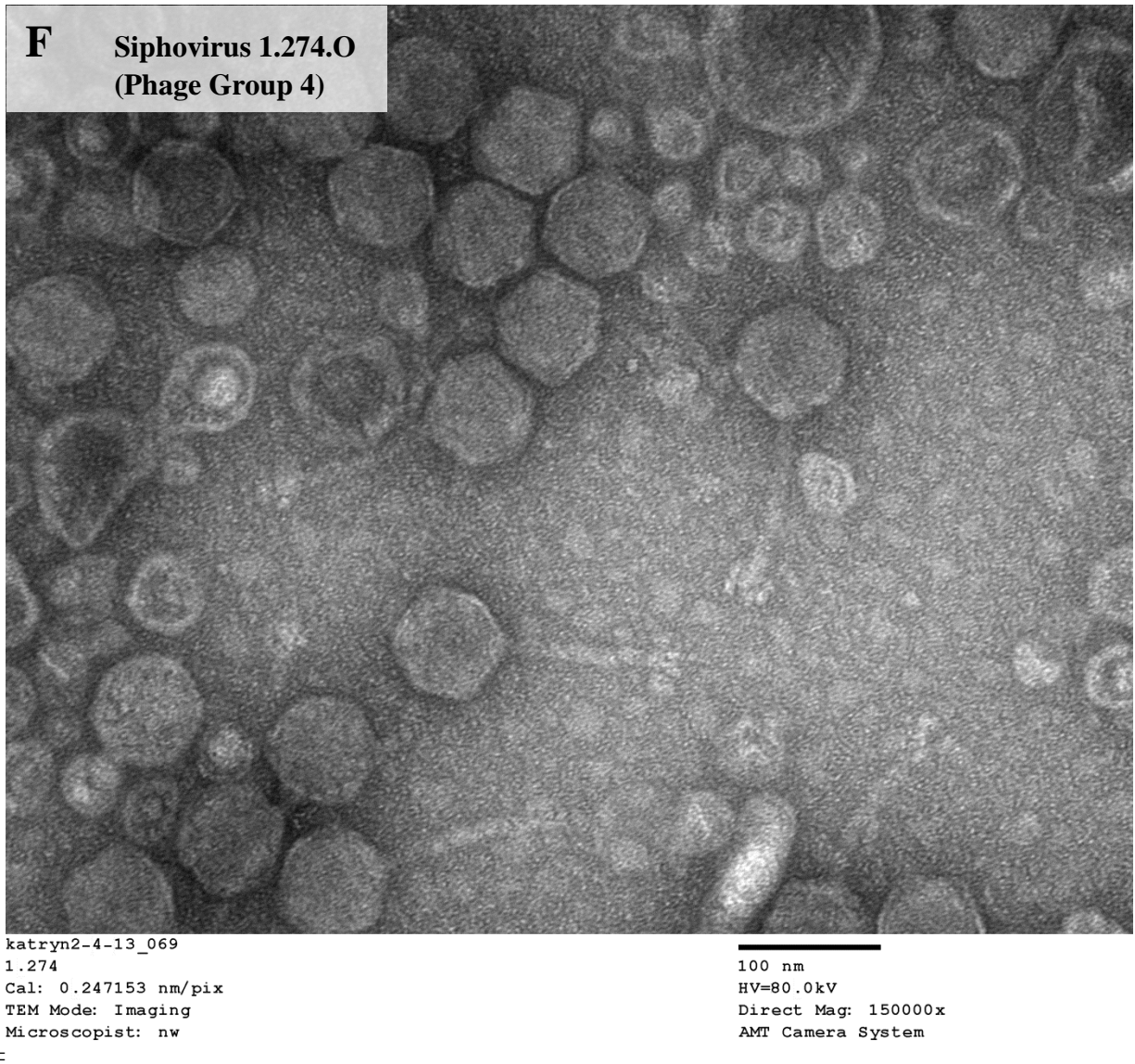


Figure 5 (Panel 5 of 9). Electron microscopy of phages. Electron microscopy by Nicki Watson.

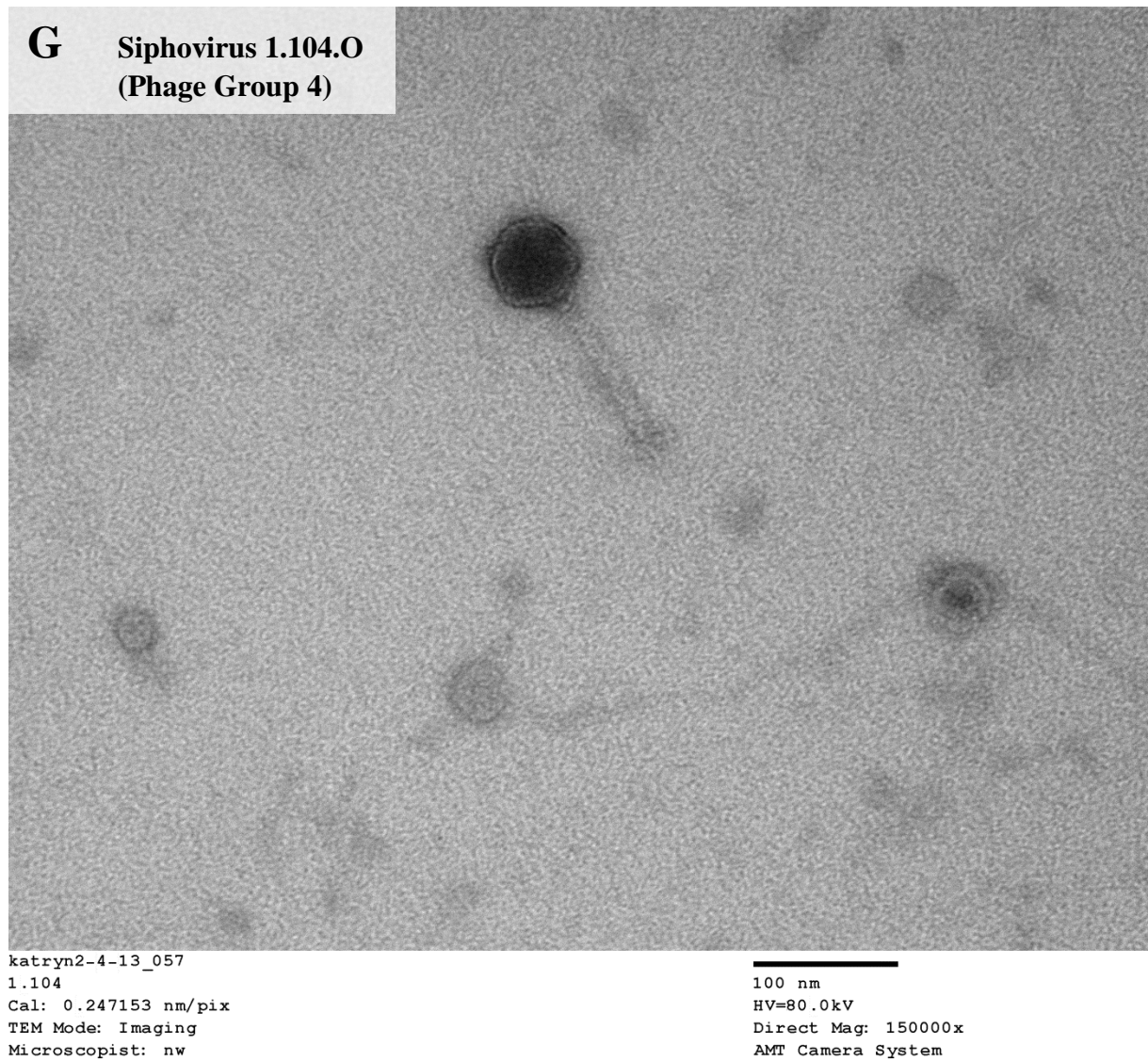
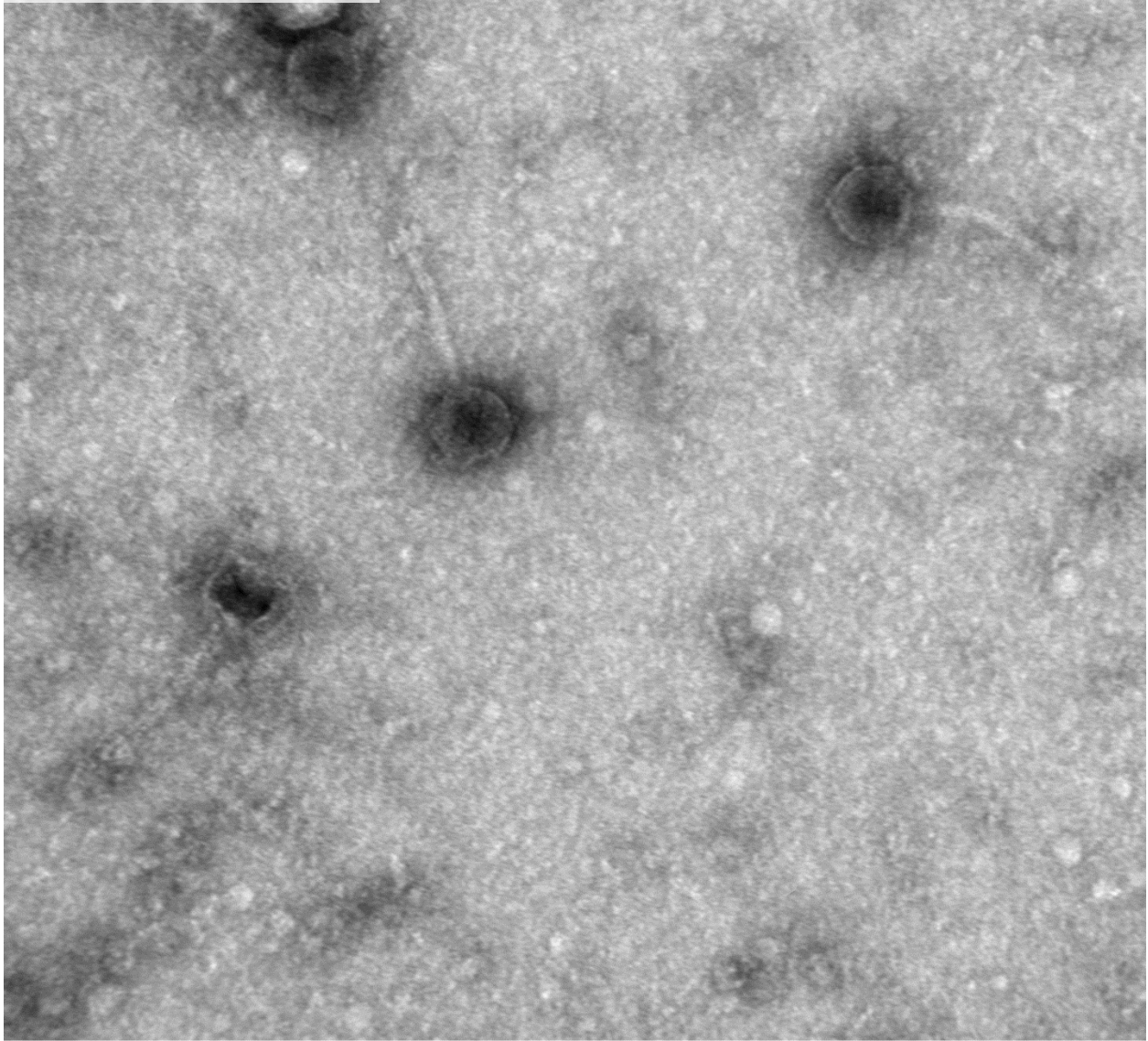


Figure 5 (Panel 6 of 9). Electron microscopy of phages. Electron microscopy by Nicki Watson.

H

**Siphovirus 1.074.O
(Phage Group 6)**



1.tif

1.74.A

Print Mag: 238000x @ 7.5 in

10:37:39 a 02/04/14

100 nm

HV=60.0kV

Direct Mag: 200000x

Figure 5 (Panel 7 of 9). Electron microscopy of phages. Electron microscopy by C. Haase-Pettingell.

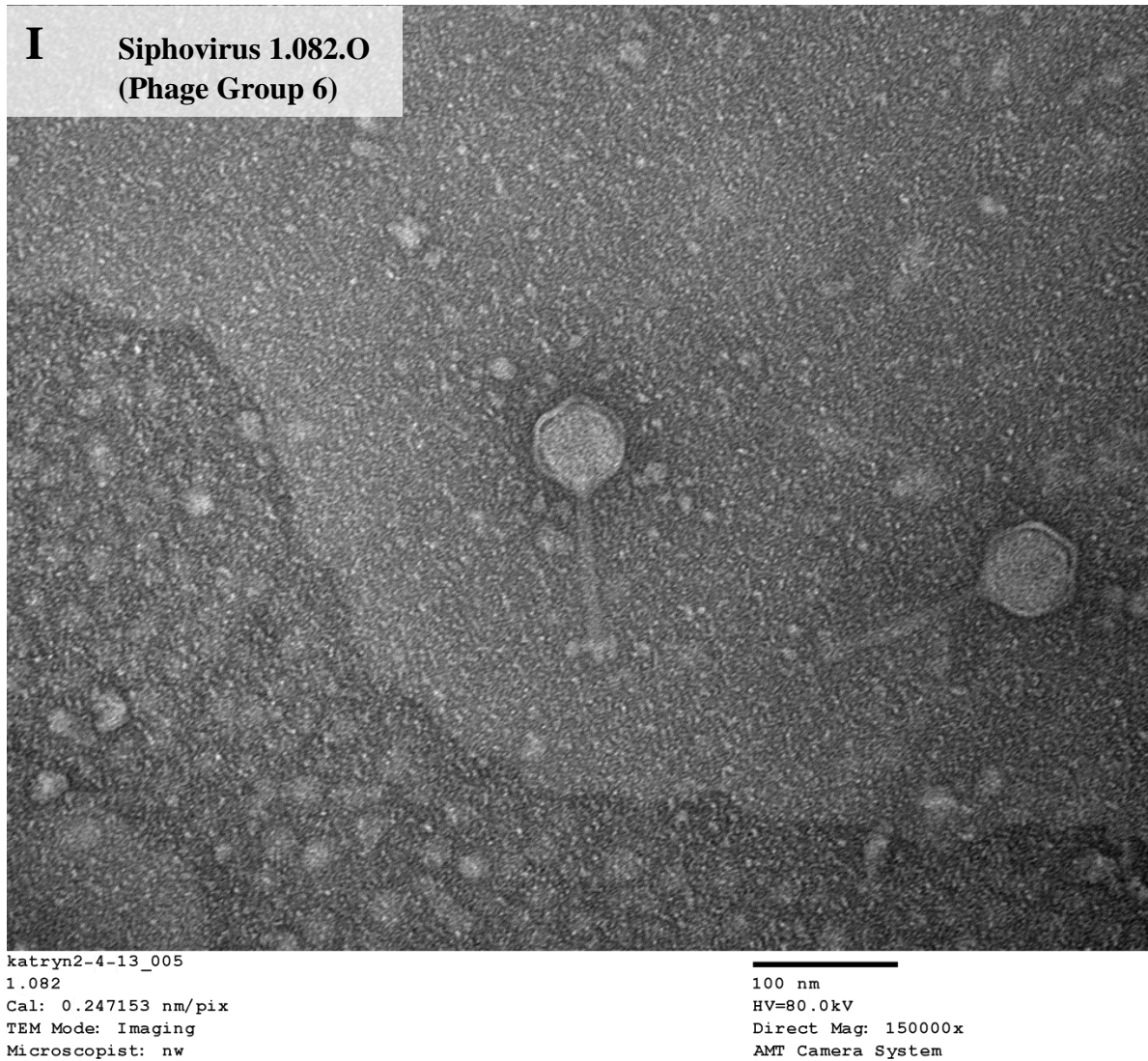


Figure 5 (Panel 8 of 9). Electron microscopy of phages. Electron microscopy by Nicki Watson.

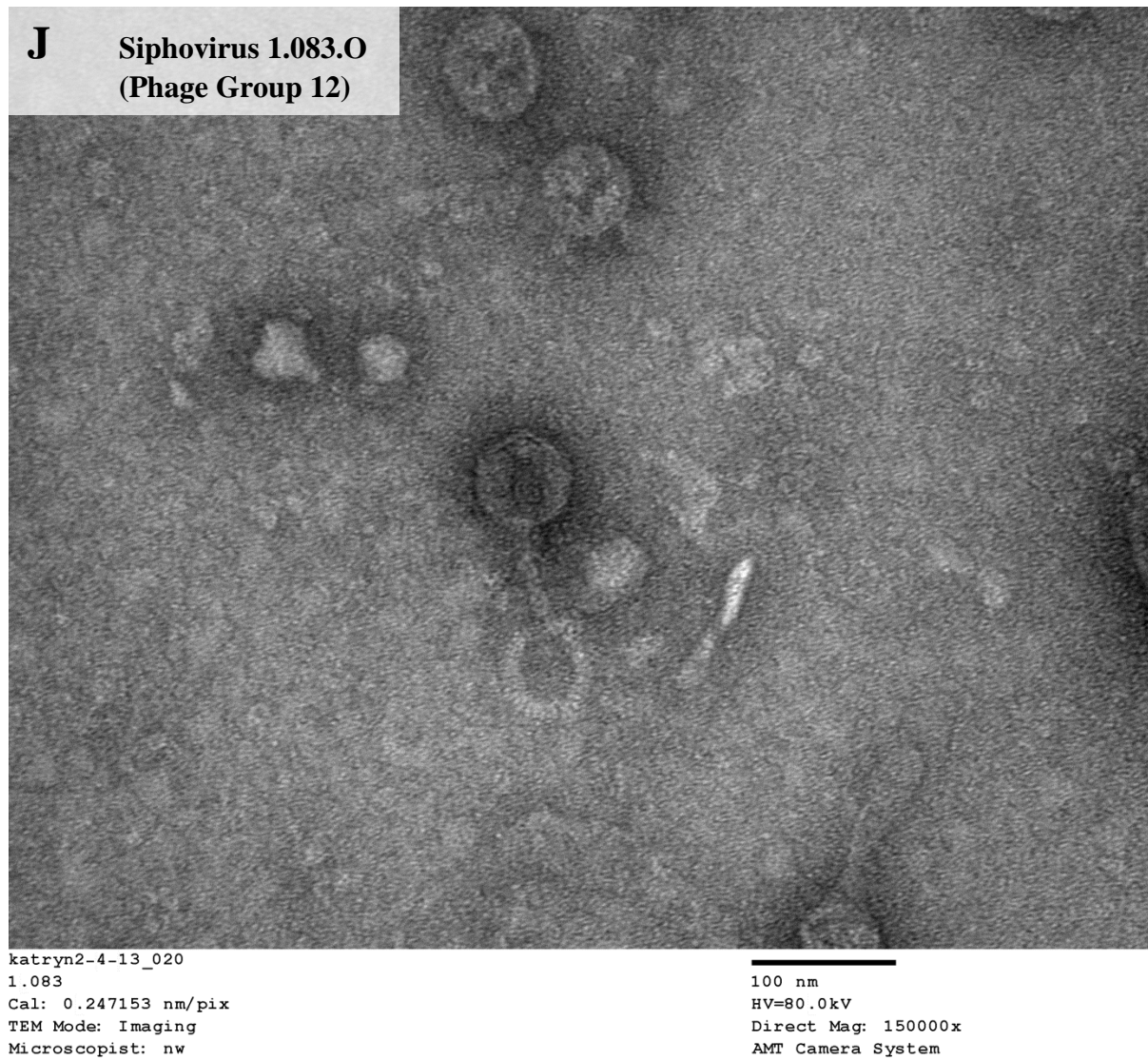
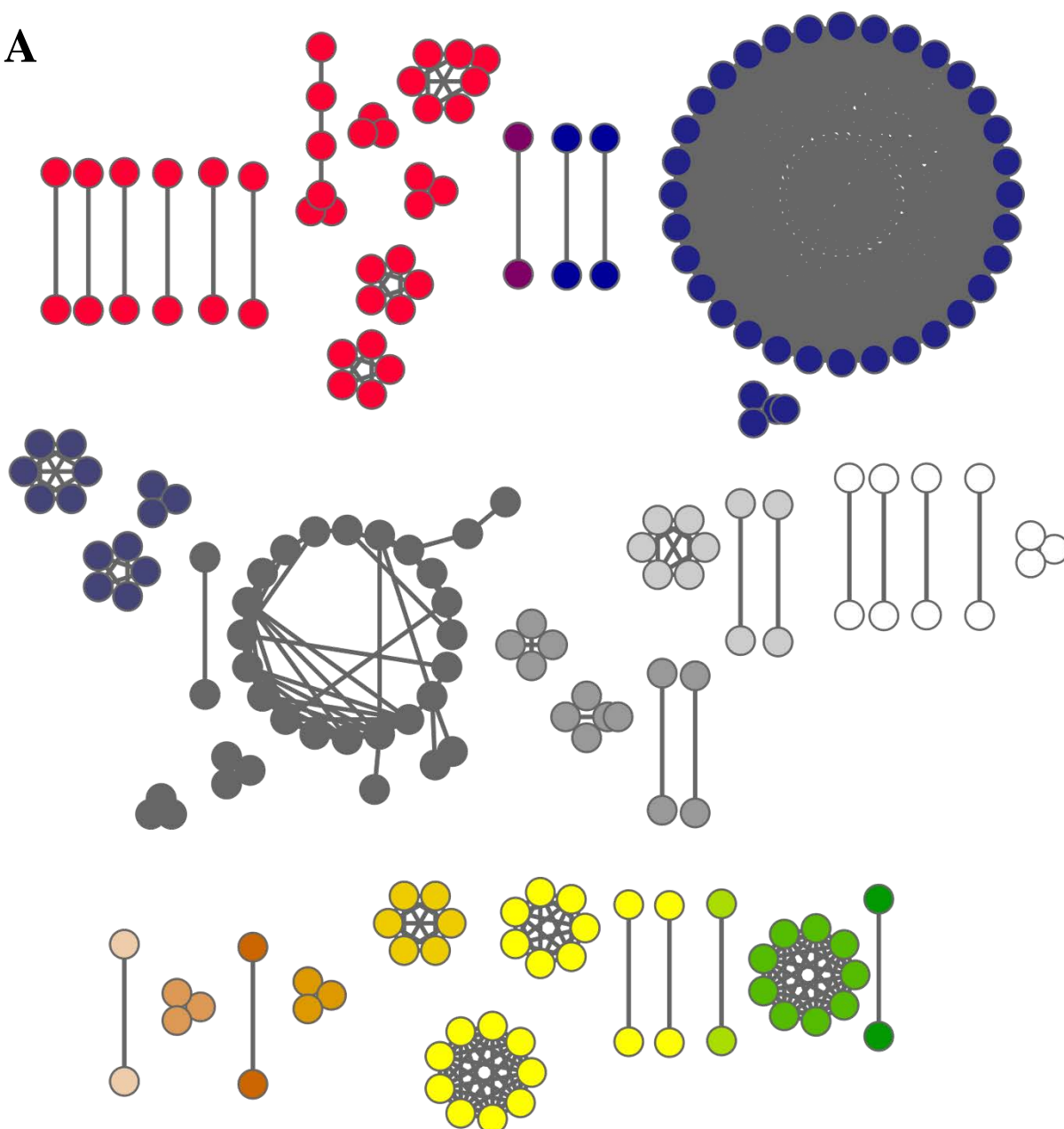


Figure 5 (Panel 9 of 9). Electron microscopy of phages. Electron microscopy by Nicki Watson.

A

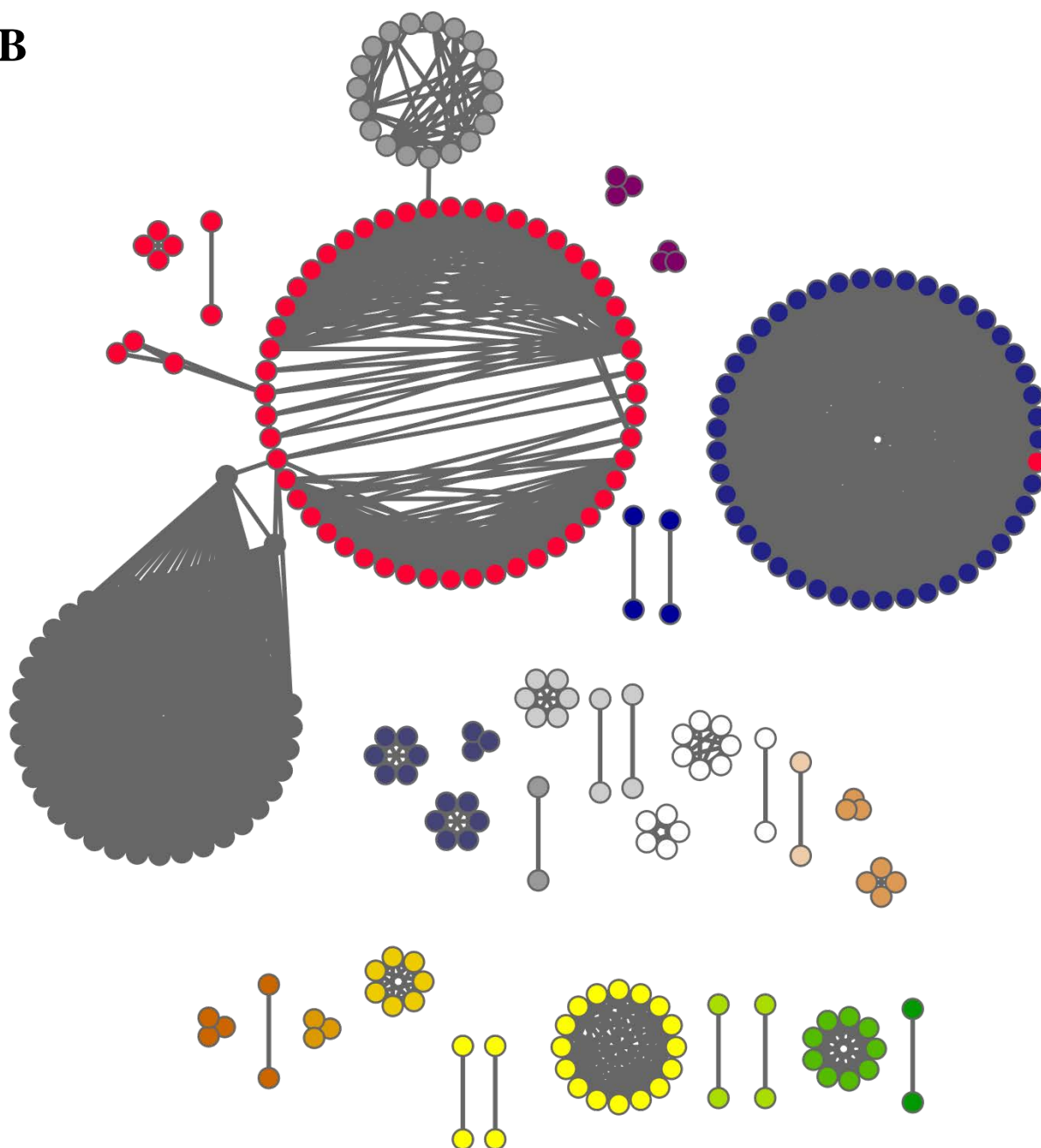


Legend: Phage genome group (G) color



Figure 6new (panel 1 of 3). Phage genomes connected by sharing at least one sequence of ≥ 500 bp at $\geq 99\%$ identity (A), $\geq 95\%$ identity (B), $\geq 75\%$ identity (C). Nodes represent phage genomes and are colored by phage genome group as indicated in legend.

B

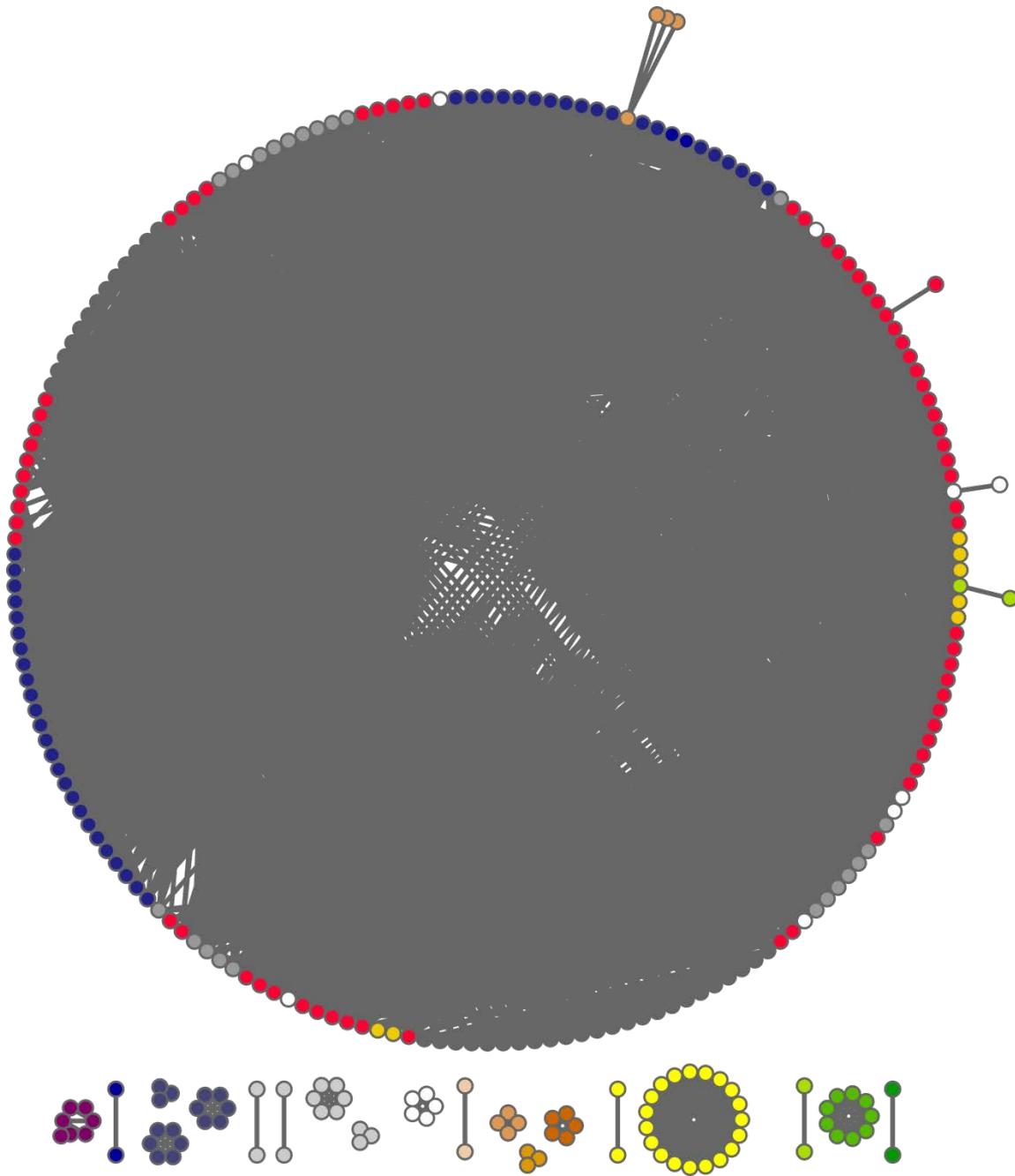


Legend: Phage genome group (G) color



Figure 6 (panel 2 of 3). Phage genomes connected by sharing at least one sequence of ≥ 500 bp at $\geq 99\%$ identity (A), $\geq 95\%$ identity (B), $\geq 75\%$ identity (C). Nodes represent phage genomes and are colored by phage genome group as indicated in legend.

C



Legend: Phage genome group (G) color



Figure 6 (panel 3 of 3). Phage genomes connected by sharing at least one sequence of ≥ 500 bp at $\geq 99\%$ identity (A), $\geq 95\%$ identity (B), $\geq 75\%$ identity (C). Nodes represent phage genomes and are colored by phage genome group as indicated in legend.

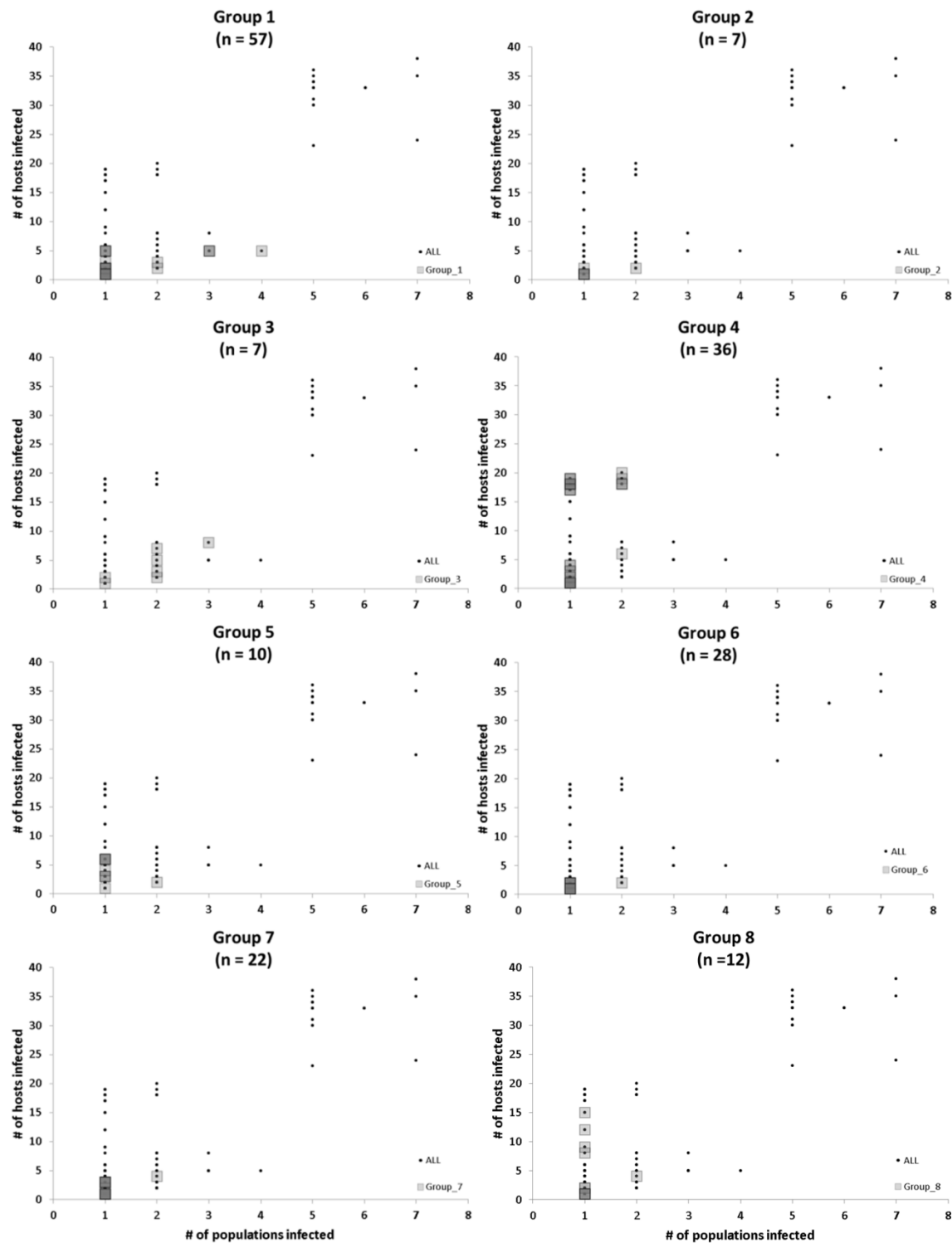


Figure 7 (1st of 3 panels). Infection profiles of phage genome groups with respect to number of host populations and total number of hosts. Infection profiles of target group highlighted in blocks over background circles indicating infections by all phages.

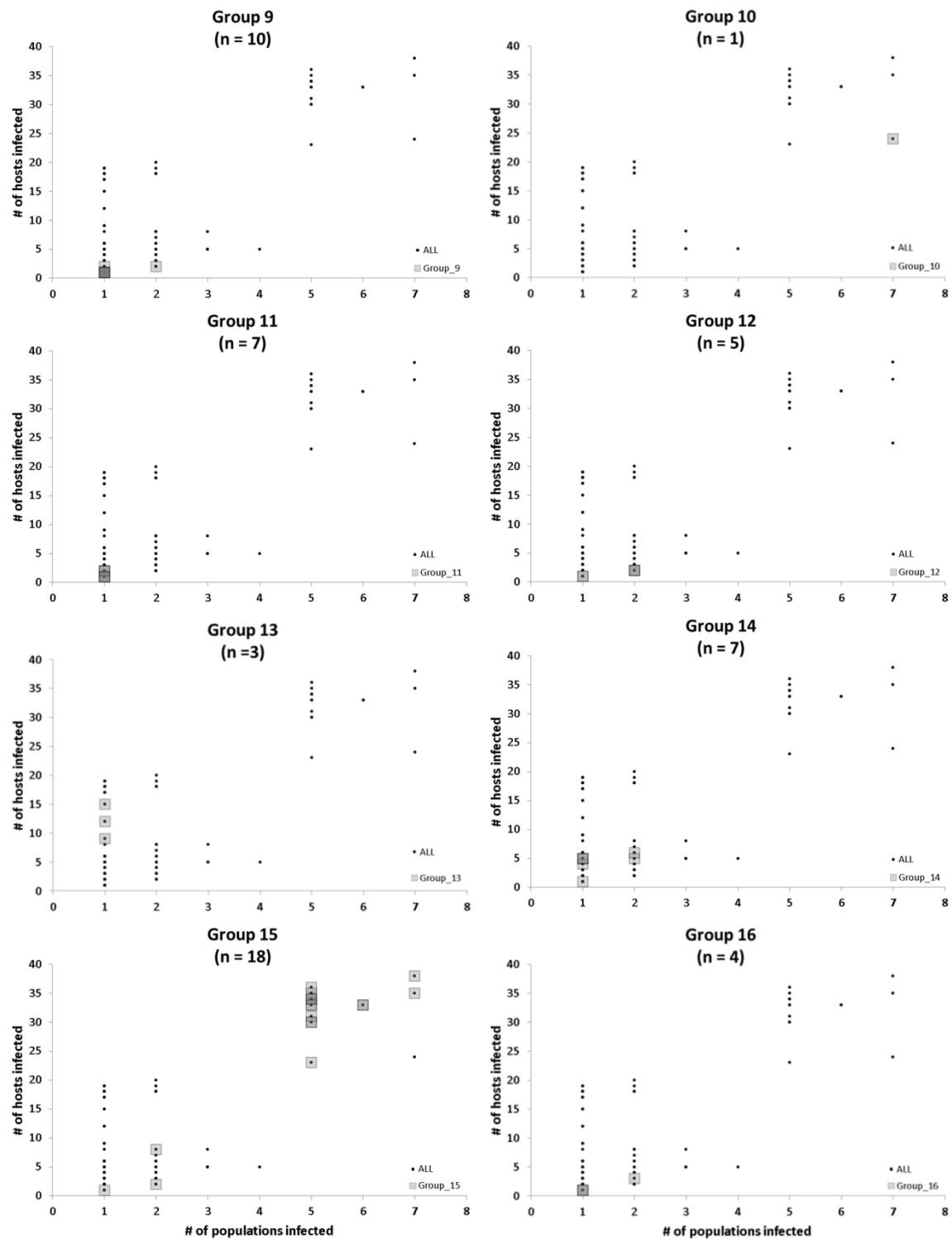


Figure 7 (2nd of 3 panels). Infection profiles of phage genome groups with respect to number of host populations and total number of hosts. Infection profiles of target group highlighted in blocks over background circles indicating infections by all phages.

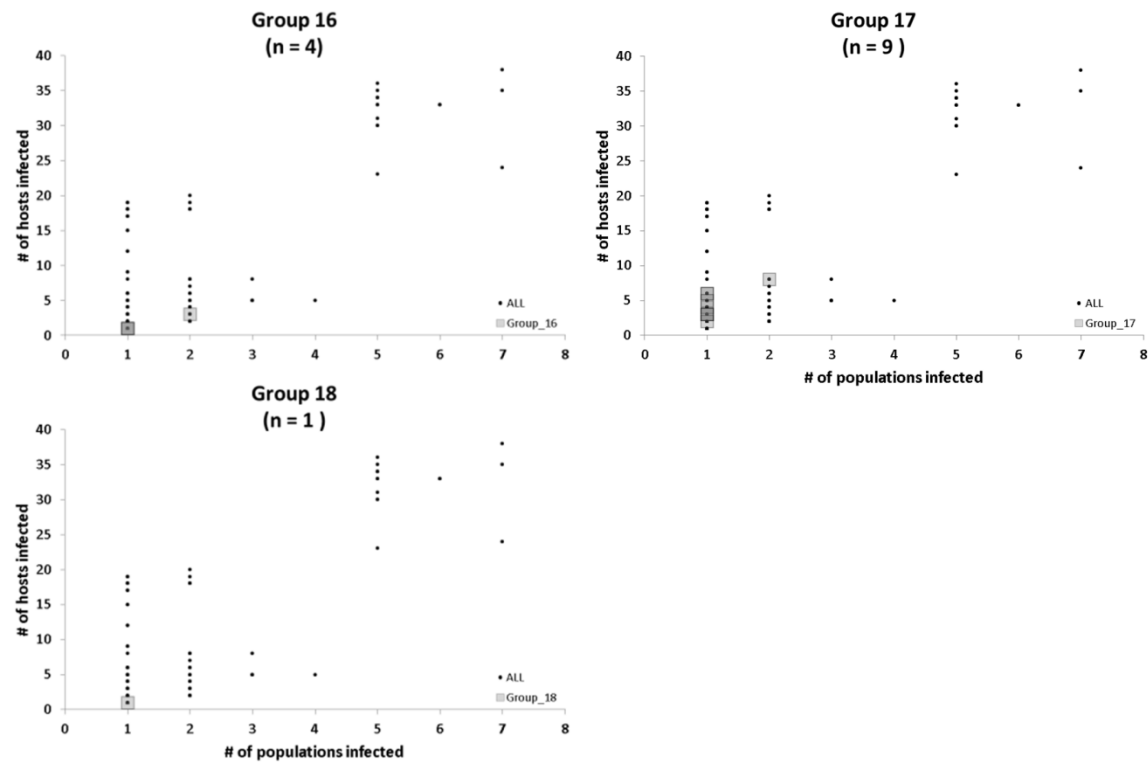


Figure 7 (3rd of 3 panels). Infection profiles of phage genome groups with respect to number of host populations and total number of hosts. Infection profiles of target group highlighted in blocks over background circles indicating infections by all phages.

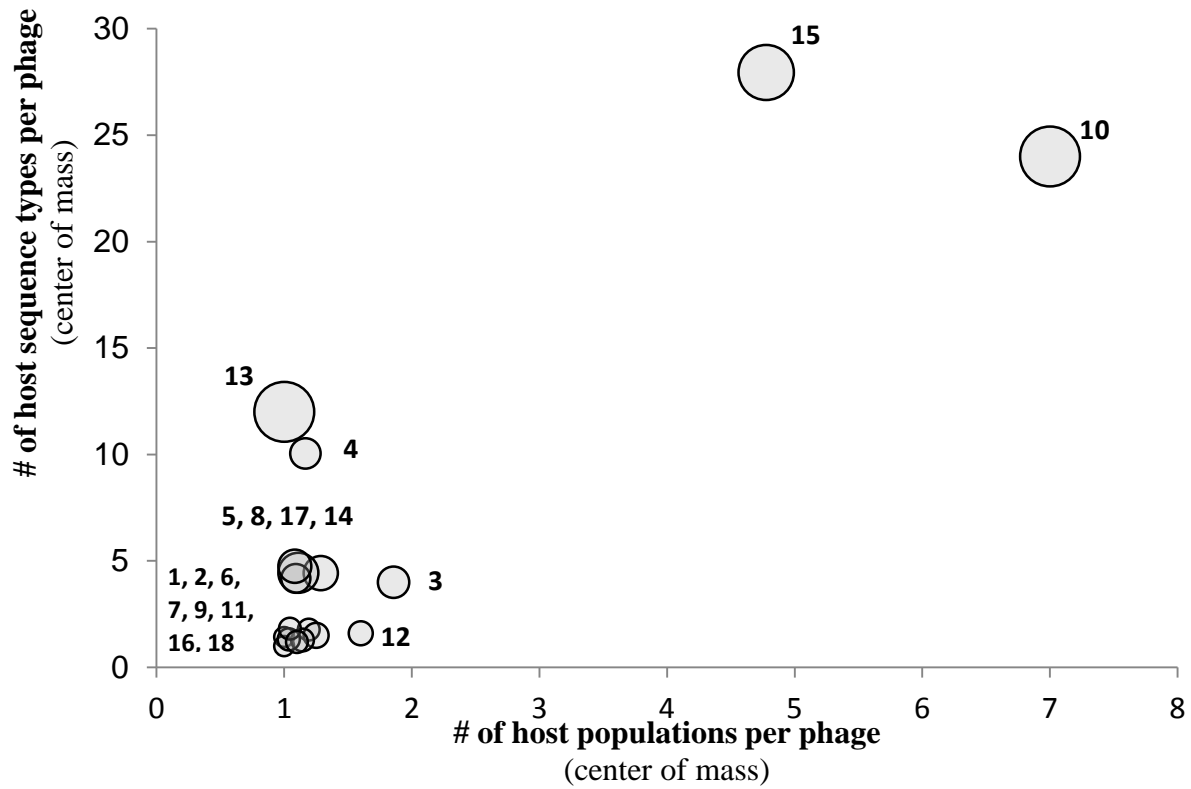


Figure 8. Phage genome group infection profiles, with respect to number of host populations and host sequence types infected. Color intensity increases with number of overlapping points($n = 18$), bubble size represents population center of mass for number of days over which hosts were infected, label is genome group number.

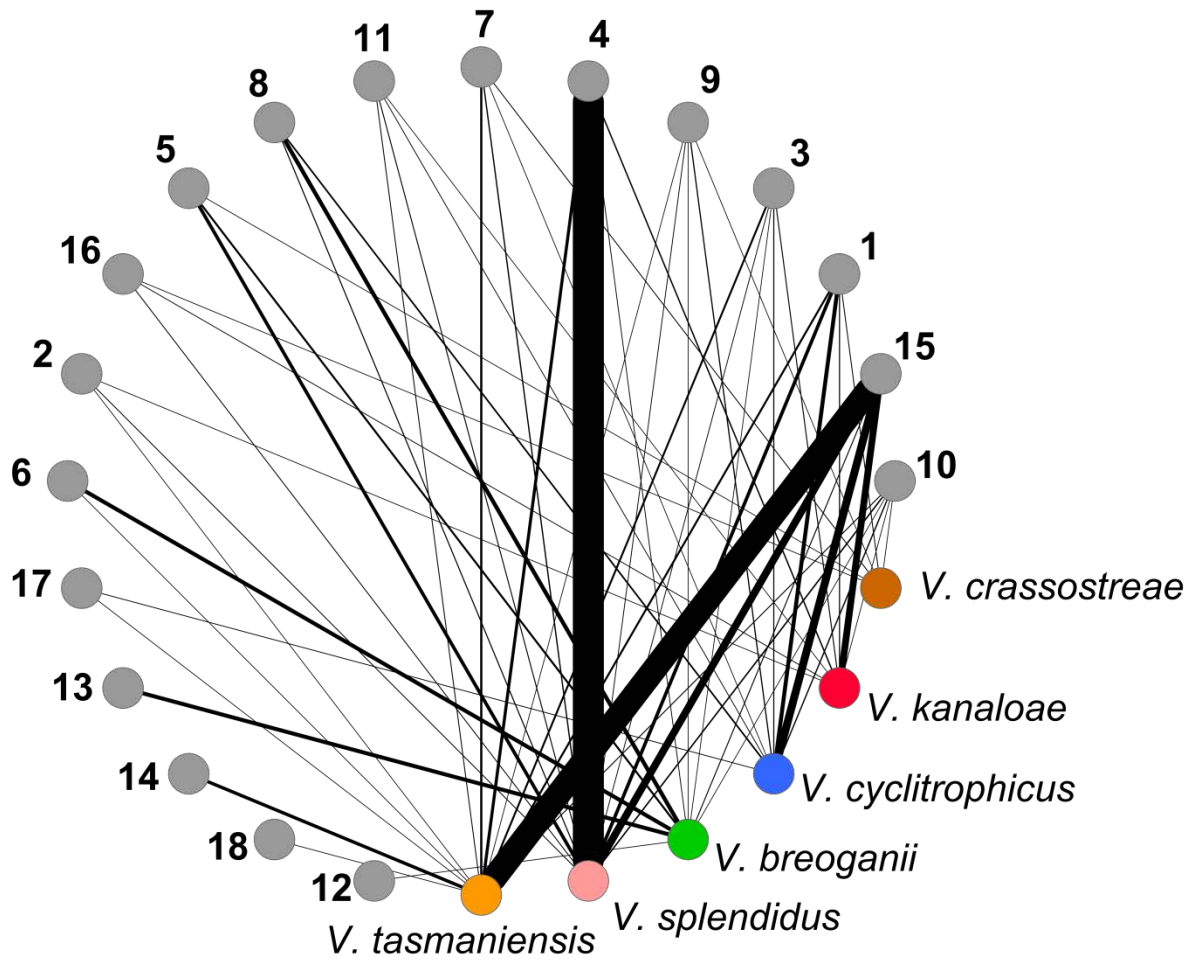


Figure 9. Number of infections per *Vibrio* population, by phage group. Degree sorted layout, line width proportional to number of interactions, phage groups indicated by numbers.

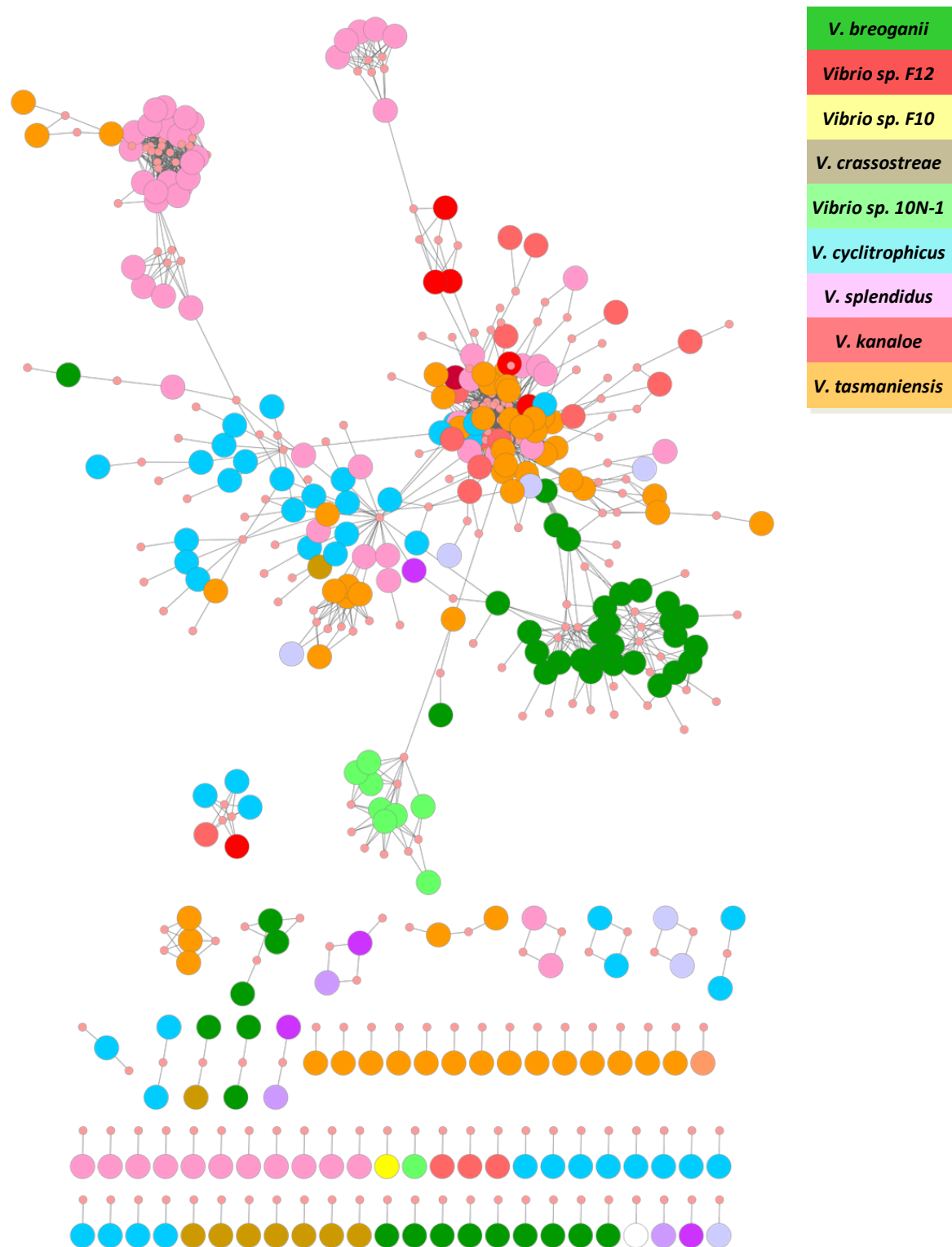


Figure 10. Network representation of host range matrix, with respect to hosts with known populations. Pink nodes represent phages, colored nodes represent hosts strains with populations indicated by different colors, as identified in legend.

Table 1. Differences in diversity, susceptibility to phages, and phage predator load across days.

	All Days	Day 222	Day 261	Day 286
Total Strains	1334	460	468	406
Susceptible Strains	295	44	109	142
Proportion Susceptible	0.22	0.1	0.23	0.35
Average PFU mL⁻¹ when positive	0.77	0.38	0.58	1.04
Proportion of Strains with > 1 PFU mL⁻¹	0.04	0.01	0.04	0.07
Total # <i>Hsp60</i> Sequence Types (STs)		444	437	390
Unique <i>Hsp60</i> STs		155	206	193
Evenness of <i>Hsp60</i> STs		0.784	0.919	0.916
Shannon Diversity Index of <i>Hsp60</i> STs		3.956	4.897	4.822

Table 2. Differences in nutrient concentrations across sampling days and over preceding days.

<i>Day</i>	uM NH₄	uM Silicate	uM PO₄	uM NO₂+NO₃
220	0.38	4.45	0.24	1.76
221	<0.05	4.11	<0.05	<0.05
222	0.22	3.99	0.10	0.32
222 - 3 day average	0.21	4.18	0.13	0.71
259	1.41	8.60	0.22	1.64
260	0.85	7.57	0.24	2.93
261	<0.05	5.47	0.07	2.11
261 - 3 day average	0.77	7.21	0.18	2.23
284	1.67	8.18	0.37	4.69
285	1.02	10.96	0.56	22.38
286	0.50	8.02	0.32	6.43
286 - 3 day average	1.06	9.06	0.42	11.17

Table 3. Differences in temperatures across sampling days and over preceding week

<i>Day</i>	Air Temperature C	Water Temperature C
216	27.0	10.9
217	24.5	10.7
218	28.7	13.5
219	24.2	11.2
220	26.3	11.0
221	26.4	12.8
222	25.6	13.8
222 – week average	26.1	12.0
255	16.5	15.2
256	18.5	16.0
257	20.8	16.1
258	18.1	16.6
259	19.7	14.9
260	15.6	15.0
261	19.0	16.3
261 – week average	18.3	15.7
280	12.4	14.7
281	16.0	15.0
282	15.1	14.9
283	15.8	14.9
284	19.3	n/a
285	13.2	14.4
286	13.5	14.2
286 – week average	15.0	14.6

Table 4. Phage susceptibility of main *Vibrio* populations. Populations abbreviated as follows: *Vibrio breoganii* (breo), *V.crassostreae* (crass), *V.cyclitrophicus* (cyc), *V.kanaloae* (kan), *V.splendidus* (splen), *V. tasmaniensis* (tas). Test for differences in proportions among populations by 6-sample test for equality of proportions without continuity correction. Significance groups for post hoc tests of pairwise comparison of proportions with Holm correction indicated by superscript in % susceptible ($p < 0.05$).

	breo	crass	cyc	kan	splen	tas	Total focal strains
Day 222	$(\chi^2 = 49.496, df = 5, p\text{-value} = 1.76e-9)$						
Total # of Strains	30	54	37	141	91	74	427
% of total focal strains	7%	13%	9%	33%	21%	17%	100%
# susceptible strains	13	6	4	3	10	5	41
% susceptible	43% ²	11% ¹	11% ¹²	2% ¹	11% ¹	7% ¹	10% ¹
Day 261	$(\chi^2 = 24.754, df = 5, p\text{-value} = 1.55.e-4)$						
Total # of Strains	68	28	63	25	86	95	365
% of total focal strains	19%	8%	17%	7%	24%	26%	100%
# susceptible strains	23	5	7	9	11	34	89
% susceptible	34% ²	18% ¹²	11% ¹	36% ¹²	13% ¹	36% ¹²	24%
Day 286	nsd						
Total # of Strains	17	17	80	19	120	68	321
% of total focal strains	5%	5%	25%	6%	37%	21%	100%
# susceptible strains	8	1	35	8	40	25	117
% susceptible	47%	6%	44%	42%	33%	37%	36%
ALL Days	$(\chi^2 = 41.717, df = 5, p\text{-value} = 6.72.e-8)$						
Total # of Strains	115	99	180	185	297	237	1113
% of total focal strains	10%	9%	16%	17%	27%	21%	100%
# susceptible strains	44	12	46	20	61	64	247
% susceptible	38% ⁴	12% ¹²	26% ²³	11% ¹	21% ¹²³	27% ³⁴	22%

Table 5. Distribution of main populations across size fractions, shown as percentage of strains. Isolates for each population are pooled from those collected across all three sampling days. Sign of the significant posthoc Pearson's χ^2 test residuals for populations indicated with positive in green and negative in red ($\chi^2 = 278.1234$, df = 15, p-value < 2.2e-16).

	0.2 uM	1 uM	5 uM	63 uM
<i>Vibrio breoganii</i>	15 %	23 %	50 %	11 %
<i>Vibrio crassostreae</i>	2 %	17 %	35 %	45 %
<i>Vibrio cyclitrophicus</i>	14 %	11 %	10 %	64 %
<i>Vibrio kanaloae</i>	38 %	38 %	18 %	6 %
<i>Vibrio splendidus</i>	26 %	26 %	24 %	25 %
<i>Vibrio tasmaniensis</i>	28 %	26 %	24 %	22 %

Table 6. Phage infection profile summaries. (n = 253, excludes interactions with hosts not mapped to populations)

	# of Phages (n = 253)	% of phages	# of infections	% of total infections
# of host strains				
1	116	46%	116	8%
1 < x < 10	93	37%	340	22%
>10	44	17%	1065	70%
# of host <i>Hsp60</i> sequence types				
1	180	71%	524	34%
1 < x < 10	51	20%	313	21%
>10	22	9%	684	45%
# of host days				
1	166	66%	339	22%
2	58	23%	422	28%
3	29	11%	760	50%
# of host populations				
1	199	78.7%	634	42%
2	27	10.7%	175	12%
3	4	1.6%	23	2%
4	2	0.8%	23	2%
5	15	5.9%	467	31%
6	2	0.8%	66	4%
7	4	1.6%	133	9%

Table 7. Properties of Nahant Collection host range matrix & comparison with Moebus & Nattkemper matrix (Flores et al., 2013; Moebus and Nattkemper, 1981). Hosts: Nahant strains that were identified as susceptible to co-isolated phages on any of three separate sampling days (direct plating). Phages: Phages isolated from each host during direct plating, generally limited to one for each susceptible host strain.

	MN Atlantic Series Matrix (Flores et al., 2013; Moebus and Nattkemper, 1981)	ALL Days	Within Day 222	Within Day 261	Within Day 286
# of Hosts (H)	286	277	43	99	135
# of Phages (P)	215	260	38	96	126
# of Interactions (I)	1332	1573	74	282	694
# of Species (S = H+P)	501	537	81	195	261
Connectance (C = I/M)	0.02	0.02	0.05	0.03	0.04
Host mean interactions ($L_H = I/H$)	4.66	5.68	1.72	2.85	5.14
Phage mean interactions ($L_I = I/P$)	6.2	6.1	1.9	2.9	5.5
Size (M = H*P)	61,490	72,020	1,634	9,504	17,010

Table 8 (1st of 2 panels). Phage group infection profiles. Only infections between sequenced phage genomes and hosts with mapped populations were included (1351 infections). Abbreviations: host sequence type (HST), center of mass (com).

Phage Group	# of phages	Total # of populations infected	Total # of HSTs infected	Total # of days isolated	# of host populations (com)	# of host strains (com)	# of host days (com)
1	57	9	36	3	1.2	1.8	1.1
2	7	4	7	2	1.1	1.3	1.1
3	7	7	16	2	1.9	4.0	1.6
4	36	4	15	3	1.2	10.1	1.5
5	11	3	8	3	1.1	4.2	1.5
6	28	2	9	3	1.0	1.3	1.1
7	22	5	17	3	1.0	1.8	1.1
8	12	4	20	3	1.1	4.8	1.7
9	10	6	11	3	1.1	1.2	1.1
10	1	7	18	1	7.0	24.0	3.0
11	7	4	7	3	1.0	1.4	1.0
12	5	3	4	1	1.6	1.6	1.2
13	3	1	6	2	1.0	12.0	3.0
14	7	2	3	2	1.3	4.4	1.7
15	18	8	37	3	4.8	27.9	2.8
16	4	3	4	2	1.3	1.5	1.3
17	9	3	3	3	1.1	4.4	2.0
18	1	1	1	1	1.0	1.0	1.0

Table 8 (2nd of 2 panels). Phage group infection profiles. Only infections between sequenced phage genomes and hosts with mapped populations were included (1351 infections). Abbreviations: host sequence type (HST), center of mass (com).

Phage Group	# of phages	Total # of Infections	% of Infections
1	57	101	7%
2	7	9	1%
3	7	28	2%
4	36	362	27%
5	11	46	3%
6	28	37	3%
7	22	40	3%
8	12	57	4%
9	10	12	1%
10	1	24	2%
11	7	10	1%
12	5	8	1%
13	3	36	3%
14	7	31	2%
15	18	503	37%
16	4	6	0%
17	9	40	3%
18	1	1	0%

Table 9 (Panel 1 of 2). Summary of expected morphotypes by phage group.

Phage Group	Expected Morphotype	Electron microscopy representative	Summary of significant BLAST hits ($e < 10^{-5}$) of representative proteins from protein clusters in each phage group against NCBI and ACLAME
1	<i>Siphoviridae</i>	Vibrio phage Jenny 12G5	Multiple hits to <i>Vibrio</i> phage Jenny 12G5 (siphovirus)
2	<i>Myoviridae</i>	Vibrio phages Helene, Eugene, & 1.084.O	Multiple hits to myoviruses (Vibrio phage Helene, Vibrio phage Eugene, Vibrio phage 1CP1-2004A)
3	unclassified	none	Multiple hits to siphoviruses (<i>Listonella</i> phage phiHSIC, <i>Vibrio</i> phage pVp-1, <i>Vibrio</i> phage Jenny 12G5, HK97) but also to odd case of VP16T which has siphovirus-like genome but looks myo-like in EM
4	<i>Siphoviridae</i>	Vibrio phages 1.104.O & 1.274.O	Multiple hits to <i>Salmonella</i> phage FSL SP-062 (siphovirus) but also to <i>Vibrio</i> phage VP16C (myovirus)
5	<i>Podoviridae</i>	none	All three 100% conserved genes in cluster are hits to <i>Vibrio</i> phage VBP47 (N4-like podovirus)
6	<i>Siphoviridae</i>	Vibrio phage 1.082.O & 1.074.O	Multiple hits to <i>Vibrio</i> phage Jenny 12G5 (siphovirus)
7	unclassified		Multiple hits <i>Vibrio</i> phages VP16T and VP16C - these phages have genomes like siphoviruses but EMs look like myoviruses (Seguritan et al., 2003)
8	<i>Podoviridae</i>		Multiple hits to podoviruses (<i>Salinivibrio</i> phage CW02, <i>Vibrio</i> phage VpV262,
9	unclassified		Sparse hits to variety of bacteria and unclassified phages (<i>Salicola</i> phage CGphi29, <i>Vibrio</i> phage PYD21-A) and myoviruses (<i>Pseudomonas aeruginosa</i> phage phiPsa374, <i>Edwardsiella</i> phage PEi21)

Table 9 (Panel 2 of 2). Summary of expected morphotypes by phage group.

Phage Group	Expected Morphotype	Electron microscopy representative	Summary of significant BLAST hits ($e < 10^{-5}$) of representative proteins from protein clusters in each phage group against NCBI and ACLAME
10	<i>Siphoviridae</i>		Many hits among conserved genes to <i>Vibrio</i> phage 1 (siphovirus)
11	unclassified		Sparse hits to variety of bacteria and phages including <i>Edwardsiella</i> phage IW-1 (podovirus), <i>Edwardsiella</i> phage PEi21 (myovirus). ACLAME hit to BcepGomr (siphovirus) and multiple putative prophages.
12	<i>Siphoviridae</i>	<i>Vibrio</i> phage 1.083.O	Multiple hits to bacteria, especially to <i>Shewanella</i> and <i>Vibrio crassostreae</i> , and phages including <i>Vibrio</i> phage VP16C (siphovirus genome, myo EM), <i>Edwardsiella</i> phage PEi21 (myovirus), <i>Salmonella</i> phage FSL SP-062 (siphovirus),
13	<i>Siphoviridae</i>		Multiple hits to bacteria and phages including many of BcepGomr (siphovirus) but also some to <i>Vibrio</i> phage VP16T (siphovirus genome, myo EM), <i>Rhizobium</i> phage RHEph04 (myovirus)
14	unclassified		Multiple hits to bacteria and phages including <i>Salmonella</i> phage FSL SP-062 (siphovirus), <i>Vibrio</i> phage VP16C (siphovirus genome, myo EM), <i>Vibrio</i> phage VpV262 (podovirus),
15	<i>Nontailed-like</i>		Multiple hits to bacteria, ACLAME hits to PR4 DNA polymerase (<i>Tectiviridae</i>) for all members of this group, hit to PM2 (<i>Corticoviridae</i>) major capsid protein P2 in all members of this group (though hits are in two separate protein clusters).
16	unclassified		Multiple hits to bacteria, especially <i>Aliivibrio</i> and <i>Vibrio parahaemolyticus</i> , and phages including <i>Vibrio</i> phage SHOU24 (unclassified), VBM1 (unclassified), <i>Vibrio</i> phage Martha.
17	<i>Myoviridae</i>		Most hits to <i>Vibrio</i> phage Martha (myovirus) and in ACLAME many hits to MuSo2 phage [a Mu-like prophage in <i>Shewanella oneidensis</i> (Heidelberg et al., 2002; Lavigne et al., 2009)].
18	unclassified		Sparse to variety of bacteria and phages including <i>Vibrio</i> phage vB_VchM-138 (myovirus), <i>Vibrio</i> phage PVA1 (podovirus), pYD38-A (unclassified),

Table 10. Phage susceptibility of main *Vibrio* populations, by phage group. Populations abbreviated as follows: *Vibrio breoganii* (breo), *V.crassostreae* (crass), *V.cyclitrophicus* (cyc), *V.kanaloae* (kan), *V.splendidus* (splen), *V. tasmaniensis* (tas).

Phage Group	# in group	breo		crass		cyc		kan		splen		tas		all
		#	%	#	%	#	%	#	%	#	%	#	%	
1	57		.	3	3	36	40	5	6	30	34	15	17	89
2	7		.		.		.	1	20	3	60	1	20	5
3	7	1	4		.	5	19	4	15	2	8	14	54	26
4	36	2	1		.		.	8	2	326	90	26	7	362
5	11	15	33	1	2		.		.	30	65		.	46
6	28	36	97		1	3	37
7	22		.	3	11	1	4		.	7	25	17	61	28
8	12	37	70		.	11	21		.	5	9		.	53
9	10	1	9	1	9	6	55		.	2	18	1	9	11
10	1	1	4	1	4	7	30	1	4	9	39	4	17	23
11	7		.		.	2	20	1	10	4	40	3	30	10
12	5	1	100		1
13	3	36	100		36
14	7		29	100	29
15	18	4	1		.	78	16	66	14	73	15	259	54	480
16	4		.	1	17		.	2	33	3	50		.	6
17	9		.		.	2	67		.		.	1	33	3
18	1		1	100	1

Chapter 5

Conclusions & Future Work

Our view of phage infection in the oceans has been undergoing continual refinement for at least 70 years. It is hypothesized that highly specific phage infection is important for maintaining the extensive microdiversity observed in environmental microbial systems (Rodriguez-Valera et al., 2009; Thingstad and Lignell, 1997; Vage et al., 2013; Våge et al., 2013; Winter et al., 2010). If this is true, then such infections are important for all the emergent properties of ecosystem functions dependent on such microdiversity. All of these properties, and the bulk viral production observed, would then be the consequence of single infections between permissive host cells and viruses that can infect them. The structure of these interaction networks is not well resolved. The aim of this thesis was to address, quantitatively, the structure of lytic phage-host interactions in the ocean, and to provide an environmentally representative model system within which to conduct future tests of hypotheses regarding the dynamics of phage infection in diverse natural communities.

It has been fairly recently that we have come to appreciate the extensive genomic microdiversity among closely related bacteria in environmental systems. Comparisons of genome size have revealed differences of up to 20% among strains that are otherwise identical at genes core to the housekeeping business of the cell (Thompson et al., 2005). Some of the genomic differences are likely to be due to the presence of large mobile elements, particularly prophages, and other of these differences are hypothesized to be due to selection for diversity in the face of lytic phage predation. It is thus not surprising that it was likely phage workers, who by essentially “phage typing” the oceans, were the first to come appreciate the extensive diversity even among apparently phenotypically identical strains (Ahrens, 1971). The culture-based study of phage infection patterns inherently requires large scales to allow for the detection of patterns in the face of the great diversity of phages and hosts and the specificity of their interactions. The studies to determine the prevalence of phage infection, and the host ranges of phages in the ocean, have often consisted of extensive samplings and large cross tests. In particular, the combined vastness of scale, and attention to biological detail, in the many studies over several decades by Karlheinz Moebus inspire awe in anyone who has done culture-based

phage work themselves. He, together with Nattkemper, conducted thousands of strain isolations and hundreds of thousands of cross tests in his time and focused on the questions we continue to work on today (Moebus, 1980; Moebus and Nattkemper, 1981). These included quantitative framing of environmental susceptibility to phages and determination of natural abundances of phages with particular specificities; investigations of the role and dynamics of lysogeny, which he came to believe was likely an important component of marine phage-host dynamics; and of consideration of these questions in light of environmentally relevant conditions (Moebus, 1983, 1991, 1992; Moebus, 1996a, b; Moebus, 1997a, b; Moebus, 1997c).

His work has provided the most quantitative framework available and one of his datasets, re-analysed in the last year, of cross infections among his Atlantic Series of phage-host systems has provided new insight into the structure of microbial interactions over multiple scales (Flores et al., 2013; Hunt et al., 2008; Moebus and Nattkemper, 1981). Unfortunately for Moebus, and for the field, molecular tools were not readily available at the time he was at work and thus all of these multitude of studies were conducted without even definitive reference to host phylogenetic relationships, much less genomes sequences for interacting members.

It is thus with a sense of great appreciation for the luxury of place and time that I have conducted these studies of the structure of phage-host interactions in natural populations. The theoretical and operational framework established for identifying discrete and ecologically differentiated populations of *Vibrio* allows the structure of phage-host interactions to be evaluated in a new and more powerful way than has ever been possible before (Cordero et al., 2012a; Cordero et al., 2012b; Hunt et al., 2008; Preheim et al., 2011a; Preheim et al., 2011b; Szabo et al., 2012). It is in light of this theoretical structure and with the operational tools developed by others to discern it, that I have evaluated patterns of infection among isolates of phages and host strains from the same natural communities. Whereas Moebus used enrichments to recover the phages during his Atlantic sampling, the methods that I use here are quantitative and meant to recover the different phages present in environmental samples in numbers representative

of their true environmental abundance. The use of enrichments in the past, though it has been necessary to yield detectable phages for most hosts in environmental systems, has blurred the relation between the character of phages recovered and their relative abundance *in situ*. Using quantitative approaches and a population framework, the studies in this thesis have revealed that ecology and population structure do yield differences in phage-host interactions, creating different regimes of susceptibility between diverse hosts and specific types of phages within the same local community. The opportunity to evaluate these patterns with respect to phage genome diversity and content has revealed that groups of phages are not equally distributed in their impacts on hosts, or genetic exchange with each other. The work being done by others to sequence the genomes of many of these hosts will provide the basis for extending these findings to the level of interactions between genomes in the context of ecology.

Like many phage workers before me I have found it necessary to develop methods for increasing the probability of recovery of phages from environmental samples. The methods presented in this thesis include ones for recovery of infective viruses from samples precipitated by the iron flocculation approach, as well as optimizations of standard techniques for increased efficiency and higher throughput. The next development of methods in this field may include microfluidic approaches for evaluating infection and lysis as these offer potential for orders of magnitude higher throughput than can be imagined with methods available today; though the development of such methods was also explored during the course of this thesis work these efforts were not fruitful and await further attention.

The methods used to study any system also imparts certain properties on it in that they operationally limit or bias the interpretations that can be made. The use of size-exclusion approaches to study viruses is nearly universal and has resulted in the disposal, on used-up pre-filters, of many phages likely to be at least equally relevant to the study phage-host interactions as those in the filtrate. Firstly, phages may be larger than filters used, but secondly, and probably far more importantly, they are associated with the particulates that are filtered out and these particulates likely represent major sites of

phage infection and development for heterotrophic bacteria in the ocean (Torrella and Morita, 1979). Future studies of phage host interactions should explicitly incorporate the role of particulate-associated infections and the role of spatial heterogeneity of structure, cell density, and nutrient concentrations, in describing the dynamics and ecology of phage infection. The newly conceived particle-based model systems designed for *in vitro* controlled studies of microbial community assembly will be especially informative in this regard. The use of reconstituted natural phage-hosts systems, as are now available as part of the Nahant Collection developed in this thesis, will allow for controlled studies of the impacts of spatial structure and natural diversity in the outcomes of phage infections.

Finally, though the work in this thesis has focused on studies of lytic interactions, the evidence for abundance of prophages provided by other work underscores that a full understanding of the interactions between viruses and hosts will require systematic study of the ecology of lysogeny (Paul, 2008). Studies of these processes will be challenging in that many of these interactions may not result in plaque formation at all, and may thus appear to be negative using standard approaches based on observation of lysis. This also underscores that the interactions described in this thesis likely represent only a percentage of the true interactions that occurred but left no mark for the eye to observe. It is hoped that the environmentally representative model system developed here will aid in the future study of these processes, as well as many others yet waiting to be revealed.

REFERENCES

- Ahrens, R., 1971. Untersuchungen zur Verbreitung von Phagen der Gattung Agrobacterium in der Ostsee. Kieler Meeresforsch 27, 102-112.
- Cordero, O.X., Ventouras, L.A., DeLong, E.F., Polz, M.F., 2012a. Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. Proc Natl Acad Sci U S A 109, 20059-20064.
- Cordero, O.X., Wildschutte, H., Kirkup, B., Proehl, S., Ngo, L., Hussain, F., Le Roux, F., Mincer, T., Polz, M.F., 2012b. Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. Science 337, 1228-1231.
- Flores, C.O., Valverde, S., Weitz, J.S., 2013. Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. The ISME Journal 7, 520-532.
- Hunt, D.E., David, L.D., Gevers, D., Preheim, S.P., Alm, E.J., Polz, M.F., 2008. Resource partitioning and sympatric differentiation among closely related bacterioplankton. Science 320, 1081-1085.
- Moebus, K., 1980. A method for the detection of bacteriophages from ocean water. Helgoländer Meeresuntersuchungen 34, 1-14.
- Moebus, K., 1983. Lytic and inhibition responses to bacteriophages among marine bacteria, with special reference to the origin of phage-host systems. Helgoländer Meeresuntersuchungen 36, 375-391.
- Moebus, K., 1991. Preliminary observations on the concentration of marine bacteriophages in the water around Helgoland. Helgoländer Meeresuntersuchungen 45, 411-422.
- Moebus, K., 1992. Further investigations on the concentration of marine bacteriophages in the water around Helgoland, with reference to the phage-host systems encountered. Helgoländer Meeresuntersuchungen 46, 275-292.
- Moebus, K., 1996a. Marine bacteriophage reproduction under nutrient-limited growth of host bacteria. I. Investigations with six phage-host systems. Marine ecology progress series. Oldendorf 14, 1-12.
- Moebus, K., 1996b. Marine bacteriophage reproduction under nutrient-limited growth of host bacteria. II. Investigations with phage-host system[H 3: H 3/1]. Marine ecology progress series. Oldendorf 14, 13-22.
- Moebus, K., 1997a. Investigations of the marine lysogenic bacterium H24. I. General description of the phage-host system. Marine ecology progress series. Oldendorf 148, 217-228.
- Moebus, K., 1997b. Investigations of the marine lysogenic bacterium H24. II. Development of pseudolysogeny in nutrient-rich broth culture. Marine ecology progress series. Oldendorf 148, 229-240.
- Moebus, K., 1997c. Investigations of the marine lysogenic bacterium H24. III. Growth of bacteria and production of phage under nutrient-limited conditions. Oceanographic Literature Review 44.
- Moebus, K., Nattkemper, H., 1981. Bacteriophage sensitivity patterns among bacteria isolated from marine waters. Helgoländer Meeresuntersuchungen 34, 375-385.

Paul, J.H., 2008. Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J* 2, 579-589.

Preheim, S.P., Boucher, Y., Wildschutte, H., David, L.A., Veneziano, D., Alm, E.J., Polz, M.F., 2011a. Metapopulation structure of Vibrionaceae among coastal marine invertebrates. *Environ Microbiol* 13, 265-275.

Preheim, S.P., Timberlake, S., Polz, M.F., 2011b. Merging taxonomy with ecological population prediction in a case study of Vibrionaceae. *Appl Environ Microbiol* 77, 7195-7206.

Rodriguez-Valera, F., Martin-Cuadrado, A.-B., Rodriguez-Brito, B., Pašić, L., Thingstad, T.F., Rohwer, F., Mira, A., 2009. Explaining microbial population genomics through phage predation. *Nature Reviews Microbiology* 7, 828-836.

Szabo, G., Preheim, S.P., Kauffman, K.M., David, L.A., Shapiro, J., Alm, E.J., Polz, M.F., 2012. Reproducibility of Vibrionaceae population structure in coastal bacterioplankton. *ISME J*.

Thingstad, T.F., Lignell, R., 1997. Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquatic Microbial Ecology* 13, 19-27.

Thompson, J.R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D.E., Benoit, J., Sarma-Rupavtarm, R., Distel, D.L., Polz, M.F., 2005. Genotypic diversity within a natural coastal bacterioplankton population. *Science* 307, 1311-1313.

Torrella, F., Morita, R.Y., 1979. Evidence by electron micrographs for a high incidence of bacteriophage particles in the waters of Yaquina Bay, Oregon: ecological and taxonomical implications. *Applied and environmental microbiology* 37, 774-778.

Vage, S., Storesund, J.E., Thingstad, T.F., 2013. Adding a cost of resistance description extends the ability of virus-host model to explain observed patterns in structure and function of pelagic microbial communities. *Environ Microbiol* 15, 1842-1852.

Våge, S., Storesund, J.E., Thingstad, T.F., 2013. SAR11 viruses and defensive host strains. *Nature* 499, E3-E4.

Winter, C., Bouvier, T., Weinbauer, M.G., Thingstad, T.F., 2010. Trade-Offs between Competition and Defense Specialists among Unicellular Planktonic Organisms: the "Killing the Winner" Hypothesis Revisited. *Microbiology and Molecular Biology Reviews* 74, 42-57.

Appendix 1

A simple and Efficient Method for Concentration of Ocean Viruses by Chemical Flocculation

A simple and efficient method for concentration of ocean viruses by chemical flocculation

Seth G. John,^{1,2*} Carolina B. Mendez,^{1,3} Li Deng,⁴ Bonnie Poulos,⁴ Anne Kathryn M. Kauffman,⁵ Suzanne Kern,^{6,7} Jennifer Brum,⁴ Martin F. Polz,⁷ Edward A. Boyle¹ and Matthew B. Sullivan^{4**}

¹Departments of ¹Earth, Atmospheric, and Planetary Sciences and ²Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA.

³Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA.

⁴Civil Architectural and Environmental Engineering, University of Texas at Austin, Austin, TX, USA.

⁵Ecology and Evolutionary Biology Department, University of Arizona, Tucson, AZ, USA.

⁶MIT/Woods Hole Oceanographic Institution Joint Program in Biological Oceanography, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

⁷Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA.

Summary

Ocean viruses alter ecosystems through host mortality, horizontal gene transfer and by facilitating remineralization of limiting nutrients. However, the study of wild viral populations is limited by inefficient and unreliable concentration techniques. Here, we develop a new technique to recover viruses from natural waters using iron-based flocculation and large-pore-size filtration, followed by resuspension of virus-containing precipitates in a pH 6 buffer. Recovered viruses are amenable to gene sequencing, and a variable proportion of phages, depending upon the phage, retain their infectivity when recovered. This Fe-based virus flocculation, filtration and resuspension method (FFR) is efficient (> 90% recovery), reliable, inexpensive and adaptable to many aspects of marine viral ecology and genomics research.

Introduction

Twenty years since the discovery that viruses are abundant in aquatic systems (Bergh *et al.*, 1989; Proctor and Fuhrman, 1990), it is now clear that they are significant ecosystem drivers through their impact on their globally important microbial hosts (Fuhrman, 1999; 2000; Weinbauer and Rassoulzadegan, 2004; Suttle, 2005; 2007; Breitbart *et al.*, 2007). For example, viral lysis of cells, which can account for a large percentage of microbial mortality, influences community composition and provides a source of organic substrate through the release of cellular contents. Further, ocean viruses transfer genes from one host cell to another via transduction (Paul, 1999), impacting the evolution of both host and phage. Perhaps most well studied, for example, the cyanobacterial viruses encode and express core photosynthesis genes obtained from their hosts (Lindell *et al.*, 2004; 2005; 2007; Millard *et al.*, 2004; Clokic *et al.*, 2006; Sullivan *et al.*, 2006; Bragg and Chisholm, 2008; Hellweger, 2009). Cyanobacterial viruses often contain other genes likely critical in ocean systems, including those involved in scavenging phosphate (Sullivan *et al.*, 2005; 2010; Weigle *et al.*, 2007; Millard *et al.*, 2009) and even nitrogen (Sullivan *et al.*, 2010) from seawater.

Investigations of wild viral populations often depend on the concentration of large volumes of water for various assays. On the one hand, less abundant viruses can often be isolated or observed only through the use of concentrated seawater samples (Seeley and Primrose, 1979). On the other hand, the expanding field of viral metagenomics requires large-scale concentrations of seawater (10s to 100s of litres) to obtain enough genetic material for sequencing (e.g. Angly *et al.*, 2006). In spite of the importance of research on wild viral populations and their dependence on concentration methods, existing large-scale concentration methods are inefficient, costly and variably reliable.

While it is possible to collect viruses from natural waters using impact filtration onto $\leq 0.02 \mu\text{m}$ pore-size filters (Steward and Culley, 2010), the low filtration speed and rapid clogging of these filters render this approach only useful for filtering smaller sample volumes (up to a few litres in oligotrophic waters). Several techniques for concentrating aquatic viruses from larger volumes have been developed, including adsorption-elution methods using

Received 23 April, 2010; accepted 7 July, 2010. For correspondence: *E-mail: sjohn@gps.caltech.edu; Tel. (+1) 626 395 2936; Fax (+1) 626 683-0621; **E-mail: mbsulli@email.arizona.edu; Tel. (+1) 520 626 6297; Fax (+1) 520-621-9903.

Re-use of this article is permitted in accordance with the Terms and Conditions set out at http://www.wileyonlinelibrary.com/onlineopen/OnlineOpen_Terms

© 2010 Society for Applied Microbiology and Blackwell Publishing Ltd

larger pore-size filters (Borrego *et al.*, 1991; Katayama *et al.*, 2002; Kamata and Suzuki, 2003) and pelleting of viruses with ultracentrifugation (Colombet *et al.*, 2007). However, these methods have drawbacks including selective adsorption of viruses to treated filters (Percival *et al.*, 2004), limited volume capacity and lack of mobility of ultracentrifugation equipment, and low or variable recoveries of viruses (Fuhrman *et al.*, 2005; Colombet *et al.*, 2007). These limitations have contributed to an increased usage of ultrafiltration methods to concentrate aquatic viruses, such as vortex flow filtration (VFF) (Paul *et al.*, 1991) and subsequently tangential flow filtration (TFF) (Wommack *et al.*, 2010).

Tangential flow filtration has been the most prominent method used to concentrate viruses from natural waters because it reduces filter clogging and allows concentration of viruses from the hundreds of litres of sample that are often necessary for genomic and metagenomic analyses of aquatic viral populations (Wommack *et al.*, 2010). While TFF is currently the most efficient means of concentrating large volumes of aquatic viruses, it requires expensive equipment (hundreds to thousands of US dollars) and several hours of processing time, and results in highly variable recoveries (2–98%) of viruses (Colombet *et al.*, 2007; Schoenfeld *et al.*, 2008), depending on factors, such as sample composition, type of TFF used, the amount of backpressure used and the operator's skill in using sample recovery techniques for backflushing of the ultrafiltration membrane. Further, these backflushing procedures render some types (e.g. Helicon spiral TFF cartridges) of these \$1000 filters unusable after approximately half a dozen uses (L. Proctor and F. Rohwer, pers. comm.). Considering limitations of the available viral con-

centration methods, we sought to develop a technique that efficiently and reliably concentrates aquatic viruses and also requires less expensive equipment, requires very little technical expertise, and can be applied under field conditions such as those encountered on oceanographic research cruises.

Here we focus on chemical techniques to develop a virus concentration method suited to marine virus research applications, by adapting flocculation based wastewater treatment techniques. Iron (Chang *et al.*, 1958; Manwaring *et al.*, 1971; Zhu *et al.*, 2005), aluminum (Chang *et al.*, 1958; Wallis and Melnick, 1967; Chaudhuri and Engelbrecht, 1970) and polyelectrolytes (Johnson *et al.*, 1967) have been used to efficiently flocculate and remove viruses from wastewater (> 99% removal). We explore a flocculation, filtration and resuspension (FFR) method using FeCl_3 as an efficient, inexpensive and non-toxic flocculant, and the use of biologically benign solvents to redissolve the iron-virus flocculate.

Results and discussion

Optimizing a chemistry-based method for recovery of ocean viruses

As expected, given success with freshwater systems (e.g. Chang *et al.*, 1958; Manwaring *et al.*, 1971; Zhu *et al.*, 2005), FeCl_3 addition led to efficient virus flocculation with very little virus in the filtrate after the addition of 1 mg Fe l^{-1} to Biosphere 2 Ocean viral-fraction seawater, and effective virus recovery from polycarbonate membrane filters (Fig. 1A). For other membrane materials, the amount of virus in the filtrate was < 10%, suggesting that

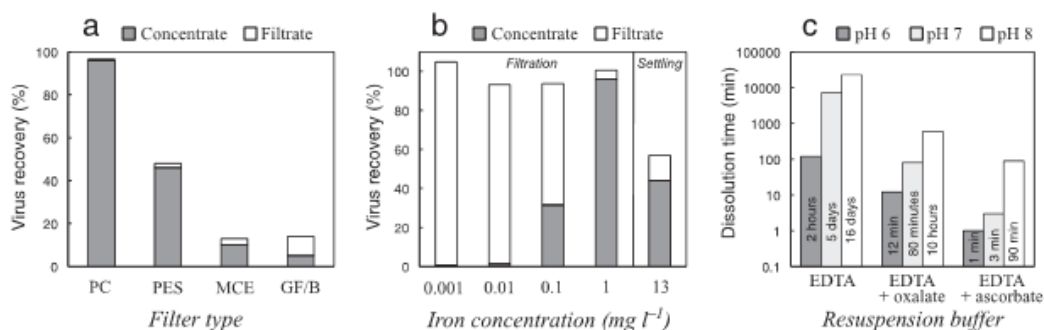


Fig. 1. Optimization of virus concentration and redissolution from Biosphere 2 Ocean viral-fraction seawater.

A. The effect of various filters on Fe-virus concentrate recovery after flocculation with 1 mg l^{-1} Fe: PC = 0.8 μm polycarbonate filters (Whatman Nuclepore), PES = 0.8 μm polyethersulfone (Pall Supor), MCE = 1.2 μm mixed cellulose ester (Millipore RAWP), and GF/B = 1.0 μm nominal pore size glass fibre filters (Whatman).

B. The effect of Fe addition on Fe-virus concentrate recovery by filtration onto a polycarbonate membrane or settling.

C. The effect of pH and resuspension buffer on the time required for dissolution of the iron hydroxide flocculate. Resuspension buffers were tested with 0.2 M EDTA in all solutions and the addition of either 0.1 M ascorbate or 0.1 M oxalate to two treatments. One millilitre of buffer was used to dissolve 1 mg of Fe.

viruses were minimally lost through the filter and rather were inadequately resuspended off the lower-yielding filter types. Optimal recovery (> 90%) was observed for Fe additions of 1 mg Fe l⁻¹ and filtration (Fig. 1B). While settling is possible in a laboratory, it is both impractical on a moving ocean research vessel and inefficient in recovering the Fe-virus precipitate even when larger amounts of Fe (e.g. 13 mg l⁻¹) are added (Fig. 1B).

Having successfully collected Fe-virus precipitate onto a polycarbonate filter, we next optimized resuspension methods to maximize recovery off the filter. To this end, we adapted marine biogeochemistry methods previously used to gently redissolve iron hydroxide precipitates while minimizing harm to phytoplankton cells during nutrient physiology studies (Tovar-Sanchez *et al.*, 2003; Tang and Morel, 2006). These techniques use a two-component mixture where the first component promotes dissolution of solid iron hydroxides and a second component chelates the Fe(III) in solution to prevent re-precipitation. Ascorbate (Anderson and Morel, 1982) and oxalate (Tovar-Sanchez *et al.*, 2003) have both been used in conjunction with ethylenediaminetetraacetate (EDTA) chelation. Ascorbate promotes iron hydroxide dissolution by reducing seawater-precipitated Fe(III) to seawater-soluble Fe(II), which can then be stabilized with EDTA chelation. Oxalate is thought to promote iron hydroxide dissolution by directly binding and liberating Fe(III) from the surface of iron hydroxide solids, releasing Fe(III) ions into solution where they can be EDTA-chelated (Cheah *et al.*, 2003; Tang and Morel, 2006). Because EDTA can inactivate viruses by binding magnesium ions (Mg²⁺) (Wells and Sisler, 1969), we provide Mg²⁺ in excess of EDTA's chelating capacity. The dissolution rate of iron hydroxide precipitate was strongly pH dependent regardless of the resuspension buffer used, with dissolution rates at pH 6 roughly two orders of magnitude greater than at pH 8. Both ascorbate- and oxalate-containing buffers acted in a more experimentally practical time frame than EDTA alone (Fig. 1C).

Based on these results, we propose the following new FeFR method for concentration of marine viruses. Seawater may first be pre-filtered (0.22 µm) to remove unicellular algae and other particulate material, depending on the needs of the researcher. One millilitre of a Fe solution (10 g FeCl₃ l⁻¹; Table 1) was added for each 10 l of viral-fraction seawater (final concentration of 1 mg Fe l⁻¹ of seawater), gently mixed and incubated 1 h at room temperature to allow Fe-virus flocculate formation. Flocculate can then be collected on a filter (142 mm diameter, 0.8 µm pore-size Whatman polycarbonate membrane filter) minimizing the overpressure (< 15 psi). Filtration time (1–2.5 h per 20 l of seawater) depends upon the sample, and filters should be replaced as needed to maintain flow rate. Place up to three filters into a 50 ml centrifuge tube, and store dark at 4°C until resuspension. Resuspend viruses by

Table 1. Solution recipes.

Concentrated Fe stock (10 g l ⁻¹ Fe):
4.83 g FeCl ₃ ·6H ₂ O into 100 ml H ₂ O
This solution is acidic and should be handled with care.
The solution has expired if a cloudy precipitate forms, do not use. Iron hydroxide precipitate will form quickly if the solution is diluted.
Ascorbate-EDTA buffer:
10 ml 2 M Mg ₂ EDTA
10 ml 2.5 M Tris HCl
25 ml 1 M ascorbic acid
Mix components and adjust to pH 6 with ~1.3 ml 10 M NaOH.
Bring to final volume of 100 ml.
A precipitate may form before the solution pH is adjusted.
Note that solution degrades quickly and should be stored in the dark at 4°C, and used within two days.
Oxalate-EDTA buffer:
10 ml 2 M Mg ₂ EDTA
10 ml 2.5 M Tris HCl
25 ml 1 M oxalic acid
Mix components and adjust to pH 6 with ~4.3 ml 10 M NaOH.
Bring to final volume of 100 ml.
A precipitate may form before the solution is adjusted to pH 6–8.
Modified SM buffer (MSM):
2.33 g NaCl
0.493 g MgSO ₄ ·7H ₂ O
5 ml 1 M Tris HCl
Mix components into ~90 ml H ₂ O.
Adjusted to pH 7.5 with 10 M NaOH.
Bring to final volume of 100 ml.
Filter sterilize.

adding 10 ml of a resuspension buffer at room temperature (0.25 M ascorbic acid, 0.2 M Mg₂EDTA, pH 6–7; Table 1), shaking occasionally by hand in order to distribute the buffer over the filters. When the precipitate has dissolved, the virus-containing buffer may be removed for subsequent processing.

Comparison of the optimized FeCl₃ method to standard methods

We compared this new viral concentration method to the standard method (TFF) using viral-fraction seawater from a Pacific Ocean viral community off of Scripps Pier in San Diego, California, USA (Fig. 2, Table 2). Four large-volume samples were concentrated using each method, including three 50 l samples and one 100 l sample by TFF and four 20 l samples by FeCl₃ flocculation. Average recoveries were 94 ± 1% (1σ SD) for FeCl₃ and 23 ± 4% (1σ SD) for TFF concentration. To confirm that the Fe-virus concentrates could be used for genetic analysis, we extracted DNA and then amplified, cloned and sequenced myovirus portal protein genes. Even with a small sample size of 10 portal protein gene sequences, the sequences obtained from the Scripps Pier Ocean water sample represented the diversity expected for a wild viral population (Fig. 3).

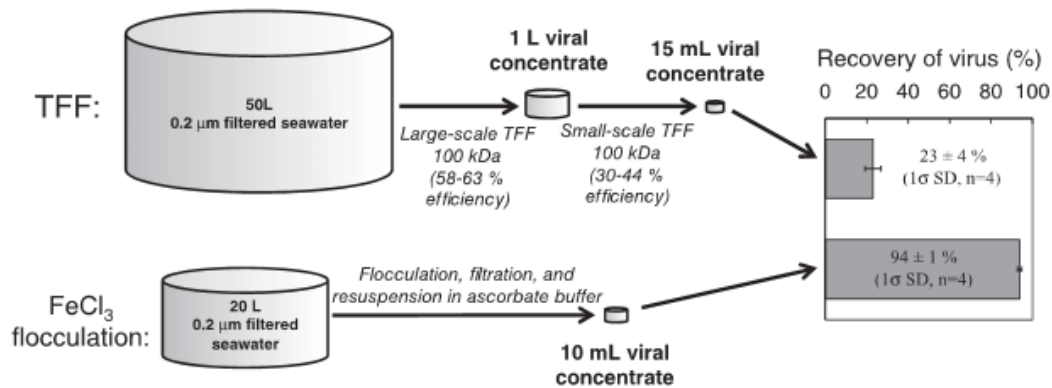


Fig. 2. Comparison of viral concentration methods showing the experimental design schematic and resulting concentration efficiency using viral-fraction (< 0.22 µm filtrate) natural seawater from Scripps Pier in San Diego, CA. Recovery is based on virus counts by epifluorescence microscopy.

Method optimization for flexible sampling needs

The method is robust to many of the variable experimental conditions that might be encountered at sea. First, incubation times of up to 12 h do not affect particle recovery, as evidenced by similar virus recoveries for four separate concentrations over 12 h (Fig. 2). Second, Fe-virus flocculate is amenable to long-term storage either with or without resuspension buffer. After 4 months of storage (dark, 4°C), 85% of virus particles were recovered from Pacific Ocean viral-fraction concentrates (data not shown). This represents only ~9% loss as compared with the initial 94% recoveries, and this was true regardless of whether the Fe-virus flocculate was resuspended immediately after filtration or 4 months later. Longer vortexing (overnight, dark, 4°C, 100 r.p.m.) of the four-month stored filters increased recovery to $92 \pm 3\%$ (1σ SD). Third, processing speeds can be halved for applications

where timing is more important than near-complete recovery of viruses. Filtering time for the same Pacific Ocean viral-fraction seawater samples was halved (25 min versus > 1 h for 20 l) by filtering the Fe-virus flocculate through a 0.22 µm pore-size polyethersulfone filter cartridge (Steripak-GP20, Millipore), with a modest drop in recovery to 71–74% ($n = 2$).

Beyond SYBR counted viral concentration efficiencies, we optimized Fe-virus concentration with a *Vibrio* phage–host system for use in culture-based studies to maximize the recovery of infective viral particles. Fe-virus concentrate resuspended in ascorbate resulted in low and unstable infective viral recovery (Fig. 4A). These poor infectivity results may be due to damage caused by free radicals formed in ascorbate (Klein, 1945; Murata *et al.*, 1986). In contrast, resuspending the Fe-virus concentrate in an oxalate buffer (0.25 M oxalate, 0.2 M Mg₂EDTA, 0.25 M Tris, pH 6; Table 1) led to efficient and stable

Table 2. Comparison of TFF and FFR viral concentration methods based on a side-by-side testing of these two methods.

	TFF	FFR
Set-up cost (\$USD)		
Prefiltration (Pump, filter holder, tubing, etc.)	~\$4000	~\$4000
Large-scale TFF	\$1603	
Small-scale TFF	\$5982	
FFR pump/filter holder		Same as prefiltration
FFR filters		\$20
Sample processing		
Volume filtered	50 l	20 l
Time for first viral concentration	1.4 ± 0.4 h (large-scale TFF)	1.4 ± 0.2 h
Time for second viral concentration	4.6 ± 0.8 h (small-scale TFF)	None needed
Time for resuspension from filter	None needed	24 h
Efficiency (% virus recovery)	$23 \pm 4\%$	$94 \pm 1\%$
Total virus recovery	1.9×10^9 viruses	3.2×10^9 viruses
Final sample volume	15 ml	10 ml

Variability in TFF methodology between labs and modifications for decreasing cost and time of FeCl₃ flocculation are discussed in the main text.

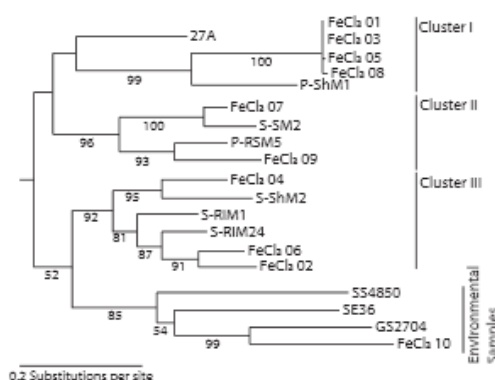


Fig. 3. Phylogenetic tree of previously published gene 20 (myovirus portal protein gene) sequences and gene sequences obtained by PCR amplification of gene 20 from Fe-virus concentrates collected at Scripps Pier. Gene sequences obtained in this study are designated as 'FeCl₃'.

infective virus recoveries (47–73% depending upon treatment) for both a myovirus and a siphovirus isolate. Infectivity was maintained through the final time points assayed when viruses were stored in oxalate or exchanged into a standard phage storage buffer (Fig. 4B–D). Additionally, infectivity was tested in a myovirus cyanophage system, S-SM1, with resuspension in either oxalate (13% recovery, $n=2$) or ascorbate (0%, $n=2$) buffer, again suggesting the choice of buffer chemistry influences infectivity. Furthermore, oxalate is advantageous because it is more stable at room temperature than buffers made with ascorbate, which must be used within 2 days of preparation (Tovar-Sanchez *et al.*, 2003).

Finally, to minimize costs, alternative set-ups might be used. For example, our samples were filtered using a peristaltic pump and a costly stainless steel 142 mm filter holder. Instead, overpressure of a seawater carboy with a home air compressor and a polycarbonate filter holder perform similarly for a total set-up cost of several hundred dollars.

Conclusions

This Fe-virus concentration method is advantageous in terms of cost, reliability and recovery efficiency. A typical TFF set-up costs over 10 thousand dollars with some costly TFF membranes having a limited lifespan. In contrast, the set-up cost for the FeCl₃ method can be as little as a few hundred dollars, with minimal per-sample costs. Further, the FeCl₃ method provides reliable, nearly complete viral recovery (92 to 95%) compared with TFF where recoveries range from 2% to 98% (Colombet *et al.*, 2007; Schoenfeld *et al.*, 2008), or $\sim 23 \pm 4\%$ as observed here.

These improvements are timely given increased sampling throughput requirements to capture temporal and spatial variability, and efforts to develop model systems from lower abundance viral types through culture-based isolations.

Experimental procedures

Wild ocean viral communities used for optimizing procedures were collected from Scripps Pier, Pacific Ocean (April 2009) and the Biosphere 2 Ocean (May 2009). Whole seawater was pre-filtered through a GF/D membrane (Whatman) in a stainless steel filter holder (Millipore, YY30-142-36) and 0.22 μ m Steripak (Millipore GP20), pressured by a peristaltic pump (MasterFlex I/P 77410-10). The 'viral fraction' seawater was subsequently concentrated using either large-scale TFF (Amersham Biosciences 100 kDa pore-size filter, UFP-100-C-9A) followed by small scale TFF (Millipore LabScale TFF System, XX42LSS11, with Pellicon XL Biomax 100 kDa pore size filter, PXB-100-C-50), or FeCl₃ flocculation and filtration using the same pump and filter holder as for the initial filtration. Virus concentrations were measured by epifluorescence microscopy after staining with SYBR Gold,

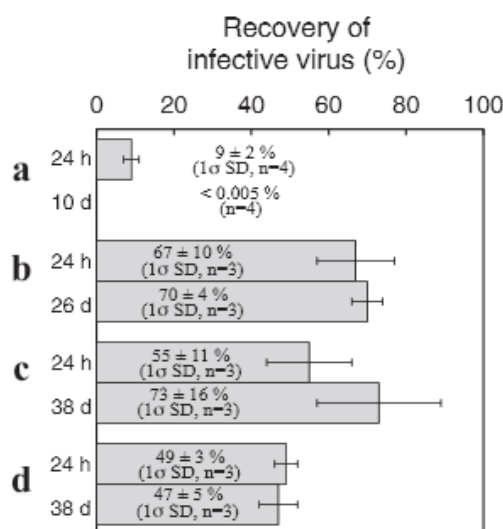


Fig. 4. Infectivity of FeCl₃-floculated viruses after variable durations of storage (24 h to 38 days) and with different resuspension buffers. Infectivity was assessed by agar overlay plaque assay of floculated and resuspended virus. Recovery was determined for (A) myovirus resuspended in ascorbate buffer, (B) myovirus resuspended in oxalate buffer and immediately transferred to modified SM buffer for long term storage, (C) myovirus resuspended in oxalate buffer, and (D) siphovirus resuspended in oxalate buffer. Viruses in (A) and (B) were spiked into artificial seawater prior to concentration, while viruses in (C) and (D) were spiked into aged natural seawater prior to concentration.

according to established procedures (Noble and Fuhrman, 1998).

The suitability of Fe-virus concentrates for genetic analyses was analysed as follows. PCR amplification of T4-like capsid assembly genes (gene 20) was obtained with primer set CPS1.1/CPS8.1 (Sullivan *et al.*, 2008) according to the following conditions: initial denaturation step of 94°C for 3 min, followed by 35 cycles of denaturation at 94°C for 15 s, annealing at 35°C for 1 min, ramping at 0.3°C s⁻¹, and elongation at 73°C for 1 min with a final elongation step at 73°C for 4 min. The PCR reactions were done in triplicate, pooled into a single tube, purified using a QIAGEN QIAquick PCR Purification kit (Qiagen, Germantown, MD, USA), cloned into a pGEM-T Easy Vector System (Promega, Madison, WI, USA) and 10 clones were then Sanger sequenced at the University of Arizona Genetics Core sequencing centre. The resulting DNA sequences were trimmed to remove PCR primers and ambiguous sequence, and aligned using Clustal X (Gap Opening penalty = 10; Extension = 0.2; DNA matrix IUB) against a suite of published gene 20 sequences chosen to represent the known diversity of these sequences in the wild (Sullivan *et al.*, 2008). The alignment was used to calculate a phylogenetic tree using PhyML under the HKY substitution model, with an empirically determined proportion of invariant sites, and transition/transversion ratio (Guindon and Gascuel, 2003).

The recovery of infective viruses from Fe-virus concentrates was tested using vibriophages and a cyanophage. The vibriophages (myovirus Vibriophage 12G01, on *Vibrio alginolyticus* 12G01; siphovirus Vibriophage Jenny 12G5, on *Vibrio splendidus* 12G5) were grown in Difco Marine Broth 2216 and spiked into 500 or 250 ml 0.2-µm-filtered seawater that lacked phages for the assayed *Vibrio* host (Kauffman, data not shown), at final concentrations of ~10⁶–10⁸ plaque-forming units (PFU) ml⁻¹. This mixture was then FeCl₃ flocculated with 4 mg Fe and filtered onto 47 mm 0.2 µm polycarbonate membranes. Replicate precipitates from separate experiments were resuspended in one of three ways: in an ascorbate buffer, in an oxalate buffer or in an oxalate buffer with subsequent transfer to modified SM (MSM) phage storage buffer (0.4 M NaCl, 0.02 M MgSO₄, 0.05 M Tris, pH 7.5) (Table 1). Transfer was achieved by centrifugal exchange with multiple rounds of centrifugation (5000 g, 20 min, room temperature) in a pre-rinsed 10K Macrosep (Pall) centrifugal device according to manufacturer's instructions, washing with ~3–4 volumes of MSM. Resuspended samples from all treatments were assayed for infective phage (PFU ml⁻¹) by agar overlay plaque assay with glycerol (Adams, 1959; Santos *et al.*, 2009). Infectivity was tested 24 h after precipitation and up to 38 days after precipitation and storage in the dark at 4°C.

The cyanophage experiments were done using similar methods, except that the cyanophage (myovirus S-SSM1, on *Synechococcus*) was grown in Pro99 medium (Moore *et al.*, 2007) and assayed for titre using the most probable number technique (Sullivan *et al.*, 2003).

Acknowledgements

Thanks to Gabriel Mitchell, Joshua Weitz, Eric Allen, Julio Ignacio, Elke Allers and Matthew Knatz for assistance in

seawater sampling, and to Julio Ignacio for generating the gene 20 tree. We acknowledge funding support from the MIT Summer Research Program (MSRP) program to CBM; the Gordon and Betty Moore Foundation to M.F.P. and Sallie W. Chisholm; NSF, and DOE to S.W.C.; the Woods Hole Center for Oceans and Human Health and Department of Energy Genomes-to-Life program to M.F.P.; Biosphere 2, BIO5, and NSF OCE0940390 to M.B.S.; and NSF EF0424599, OCE0751409 and the Arunas and Pam Chesonis Foundation to S.G.J.

References

- Adams, M.H. (1959) *Bacteriophages*. New York, USA: Interscience Publishers.
- Anderson, M.A., and Morel, F.M.M. (1982) The influence of aqueous iron chemistry on the uptake of iron by the coastal diatom *Thalassiosira weissflogii*. *Limnol Oceanogr* **27**: 789–813.
- Angly, F.E., Felts, B., Breitbart, M., Salamon, P., Edwards, R.A., Carlson, C., *et al.* (2006) The marine viromes of four oceanic regions. *PLoS Biol* **4**: 2121–2131.
- Bergh, O., Borsheim, K.Y., Bratbak, G., and Heldal, M. (1989) High abundance of viruses found in aquatic environments. *Nature* **340**: 467–468.
- Borrego, J.J., Cornax, R., Preston, D.R., Farrah, S.R., McElhaney, B., and Bitton, G. (1991) Development and application of new positively charged filters for recovery of bacteriophages from water. *Appl Environ Microbiol* **57**: 1218–1222.
- Bragg, J.G., and Chisholm, S.W. (2008) Modelling the fitness consequences of a cyanophage-encoded photosynthesis gene. *PLoS ONE* **3**: e3550.
- Breitbart, M., Thompson, L.R., Suttle, C.A., and Sullivan, M.B. (2007) Exploring the vast diversity of marine viruses. *Oceanography* **20**: 135–139.
- Chang, S.L., Stevenson, R.E., Bryant, A.R., Woodward, R.L., and Kabler, P.W. (1958) Removal of Coxsackie and bacterial viruses and the native bacteria in row Ohio River water by flocculation with aluminum sulfate and ferric chloride. *Am J Public Health* **48**: 159.
- Chaudhuri, M., and Engelbrecht, R. (1970) Removal of viruses from water by chemical coagulation and flocculation. *J Am Water Works Assoc* **62**: 563–567.
- Cheah, S.F., Kraemer, S.M., Cervini-Silva, J., and Sposito, G. (2003) Steady-state dissolution kinetics of goethite in the presence of desferrioxamine B and oxalate ligands: implications for the microbial acquisition of iron. *Chem Geol* **198**: 63–75.
- Clokier, M.R., Shan, J., Bailey, S., Jia, Y., Krisch, H.M., West, S., and Mann, N.H. (2006) Transcription of a 'photosynthetic' T4-type phage during infection of a marine cyanobacterium. *Environ Microbiol* **8**: 827–835.
- Colombet, J., Robin, A., Lavie, L., Bettarel, Y., Cauchie, H.M., and Sime-Ngando, T. (2007) Viroplankton 'pegylation': use of PEG (polyethylene glycol) to concentrate and purify viruses in pelagic ecosystems. *J Microbiol Methods* **71**: 212–219.
- Fuhrman, J.A. (1999) Marine viruses and their biogeochemical and ecological effects. *Nature* **399**: 541–548.
- Fuhrman, J.A. (2000) Impact of viruses on bacterial

- processes. In *Microbial Ecology of the Oceans*. Kirchman, D.L. (ed.). New York, NY, USA: Wiley-Liss, pp. 327–350.
- Fuhrman, J.A., Liang, X., and Noble, R.T. (2005) Rapid detection of enteroviruses in small volumes of natural waters by real-time quantitative reverse transcriptase PCR. *Appl Environ Microbiol* **71**: 4523–4530.
- Guindon, S., and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**: 696–704.
- Hellweger, F.L. (2009) Carrying photosynthesis genes increases ecological fitness of cyanophage in silico. *Environ Microbiol* **11**: 1386–1394.
- Johnson, J.H., Fields, J.E., and Darlington, W.A. (1967) Removing viruses from water by polyelectrolytes. *Nature* **213**: 665–667.
- Kamata, S.-I., and Suzuki, S. (2003) Concentration of marine birnavirus from seawater with a glass fiber filter precoated with bovine serum albumin. *Mar Biotechnol* **5**: 157–162.
- Katayama, H., Shimasaki, A., and Ohgaki, S. (2002) Development of a virus concentration method and its application to detection of enterovirus and Norwalk virus from coastal seawater. *Appl Environ Microbiol* **68**: 1033–1039.
- Klein, M. (1945) The mechanism of the virucidal action of ascorbic acid. *Science* **101**: 587–589.
- Lindell, D., Sullivan, M.B., Johnson, Z.I., Tolonen, A.C., Rohwer, F., and Chisholm, S.W. (2004) Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. *Proc Natl Acad Sci USA* **101**: 11013–11018.
- Lindell, D., Jaffe, J.D., Johnson, Z.I., Church, G.M., and Chisholm, S.W. (2005) Photosynthesis genes in marine viruses yield proteins during host infection. *Nature* **438**: 86–89.
- Lindell, D., Jaffe, J.D., Coleman, M.L., Futschik, M.E., Axmann, I.M., Rector, T., et al. (2007) Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature* **449**: 83–86.
- Manwaring, J.F., Chaudhuri, M., and Engelbrecht, R. (1971) Removal of viruses by coagulation and flocculation. *J Am Water Works Assoc* **63**: 298–300.
- Millard, A., Clokier, M.R., Shub, D.A., and Mann, N.H. (2004) Genetic organization of the psbAD region in phages infecting marine *Synechococcus* strains. *Proc Natl Acad Sci USA* **101**: 11007–11012.
- Millard, A.D., Zwirgmaier, K., Downey, M.J., Mann, N.H., and Scanlan, D.J. (2009) Comparative genomics of marine cyanomyoviruses reveals the widespread occurrence of *Synechococcus* host genes localized to a hyperplastic region: implications for mechanisms of cyanophage evolution. *Environ Microbiol* **11**: 2370–2387.
- Moore, L.R., Coe, A., Zinser, E.R., Saito, M.A., Sullivan, M.B., Lindell, D., et al. (2007) Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol Oceanogr Methods* **5**: 353–362.
- Murata, A., Suenaga, H., Hideshima, S., Tanaka, Y., and Kato, F. (1986) Virus-inactivating effect of ascorbic-acid: hydroxyl radicals as the reactive species in the inactivation of phages by ascorbic-acid. *Agric Biol Chem* **50**: 1481–1487.
- Noble, R.T., and Fuhrman, J.A. (1998) Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. *Aquat Microb Ecol* **14**: 113–118.
- Paul, J.H. (1999) Microbial gene transfer: an ecological perspective. *J Mol Microbiol Biotechnol* **1**: 45–50.
- Paul, J.H., Jiang, S.C., and Rose, J.B. (1991) Concentration of viruses and dissolved DNA from aquatic environments by vortex flow filtration. *Appl Environ Microbiol* **57**: 2197–2204.
- Percival, S.L., Chalmers, R.M., Embrey, M., Hunter, P.R., Sellwood, J., and Wyn-Jones, P. (2004) *Microbiology of Waterborne Diseases*. Amsterdam, the Netherlands: Elsevier.
- Proctor, L.M., and Fuhrman, J.A. (1990) Viral mortality of marine bacteria and cyanobacteria. *Nature* **343**: 60–62.
- Santos, S.B., Carvalho, C.M., Sillankorva, S., Nicolau, A., Ferreira, E.C., and Azeredo, J. (2009) The use of antibiotics to improve phage detection and enumeration by the double-layer agar technique. *BMC Microbiol* **9**: 10.
- Schoenfeld, T., Patterson, M., Richardson, P.M., Wommack, K.E., Young, M., and Mead, D. (2008) Assembly of viral metagenomes from Yellowstone hot springs. *Appl Environ Microbiol* **74**: 4164–4174.
- Seeley, N.D., and Primrose, S.B. (1979) Concentration of bacteriophages from natural waters. *J Appl Bacteriol* **46**: 103–116.
- Steward, G.F., and Culley, A.I. (2010) Extraction and purification of nucleic acids from viruses. In *Manual of Aquatic Viral Ecology*. Wilhelm, S.W., Weinbauer, M.G., and Suttle, C.A. (eds). Waco, TX, USA: ASLO, pp. 154–165.
- Sullivan, M.B., Waterbury, J.B., and Chisholm, S.W. (2003) Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. *Nature* **424**: 1047–1051.
- Sullivan, M.B., Coleman, M.L., Weigle, P., Rohwer, F., and Chisholm, S.W. (2005) Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLoS Biol* **3**: e144.
- Sullivan, M.B., Lindell, D., Lee, J.A., Thompson, L.R., Bielawski, J.P., and Chisholm, S.W. (2006) Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLoS Biol* **4**: e234.
- Sullivan, M.B., Coleman, M.L., Quinlan, V., Rosenkrantz, J.E., DeFrancesco, A.S., Tan, G., et al. (2008) Portal protein diversity and phage ecology. *Environ Microbiol* **10**: 2810–2823.
- Sullivan, M.B., Huang, K.H., Ignacio-Espinoza, J.C., Berlin, A.M., Kelly, L., Weigle, P.R., et al. (2010) Genomic analysis of oceanic cyanobacterial myoviruses compared to T4-like myoviruses from diverse hosts and environments. *Environ Microbiol* (in press): doi: 10.1111/j.1462-2920.2010.02280.x.
- Suttle, C.A. (2005) Viruses in the sea. *Nature* **437**: 356–361.
- Suttle, C.A. (2007) Marine viruses – major players in the global ecosystem. *Nat Rev Microbiol* **5**: 801–812.
- Tang, D.G., and Morel, F.M.M. (2006) Distinguishing between cellular and Fe-oxide-associated trace elements in phytoplankton. *Mar Chem* **98**: 18–30.
- Tovar-Sanchez, A., Sanudo-Wilhelmy, S.A., Garcia-Vargas, M., Weaver, R.S., Popels, L.C., and Hutchins, D.A. (2003) A trace metal clean reagent to remove surface-bound iron from marine phytoplankton. *Mar Chem* **82**: 91–99.
- Wallis, C., and Melnick, J.L. (1967) Concentration of viruses on aluminum phosphate and aluminum hydroxide precipi-

- tates. In *Transmission of Viruses by the Water Route*. Berg, G. (ed.). New York, USA: Interscience Publishers, pp. 129–138.
- Weigle, P.R., Pope, W.H., Pedulla, M.L., Houtz, J.M., Smith, A.L., Conway, J.F., *et al.* (2007) Genomic and structural analysis of Syn9, a cyanophage infecting marine *Prochlorococcus* and *Synechococcus*. *Environ Microbiol* **9**: 1675–1695.
- Weinbauer, M.G., and Rassoulzadegan, F. (2004) Are viruses driving microbial diversification and diversity? *Environ Microbiol* **6**: 1–11.
- Wells, J.M., and Sisler, H.D. (1969) Effect of EDTA and Mg²⁺ on infectivity and structure of southern bean mosaic virus. *Virology* **37**: 227.
- Wommack, K.E., Sime-Ngando, T., Winget, D.M., Jamindar, S., and Helton, R.R. (2010) Filtration-based methods for the collection of viral concentrates from large water samples. In *Manual of Aquatic Viral Ecology*. Wilhelm, S.W., Weinbauer, M.G., and Suttle, C.A. (eds). Waco, TX, USA: American Society of Limnology and Oceanography, pp. 110–117.
- Zhu, B., Clifford, D.A., and Chellam, S. (2005) Virus removal by iron coagulation-microfiltration. *Water Res* **39**: 5153–5161.